

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:53:30 ; Search time 12.56 Seconds

(without alignments)
562.582 Million cell updates/sec

Title: US-09-912-717-1

Perfect score: 1567
Sequence: 1 MSVFGAGQLAIFRTAAGL.....TPSSPGKLLTRSLALGCKKD 314

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1567	100.0	314	3	US-09-099-676-1
2	1567	100.0	314	4	US-09-565-910-1
3	1318	84.1	315	3	US-09-099-676-3
4	1318	84.1	315	4	US-09-565-910-3
5	294	16.8	304	1	US-08-605-150A-2
6	102	6.5	349	3	US-08-605-150A-17
7	98	6.3	367	3	US-08-605-150A-6
8	98	6.3	381	3	US-08-605-150A-4
9	97.5	6.2	543	1	US-08-375-709-17
10	97.5	6.2	543	4	US-08-752-929-17
11	97.5	6.2	543	4	US-09-090-793-10
12	95.5	6.1	1455	2	US-08-726-012B-2
13	94.5	6.0	372	3	US-08-605-150A-8
14	94.5	6.0	372	3	US-08-605-150A-10
15	91.5	5.8	1026	4	US-08-614-377A-7
16	91.5	5.8	1026	4	US-09-142-648B-7
17	88.5	5.6	722	3	US-08-961-083-84
18	88.5	5.6	722	3	US-08-605-150A-2
19	85	5.4	398	1	US-08-176-413-5
20	85	5.4	398	5	PCT-US94-14919-5
21	84	5.4	4545	2	US-08-804-227C-14
22	84	5.4	4550	2	US-08-804-227C-8
23	83.5	5.3	801	1	US-08-725-012-2
24	83	5.3	801	1	US-08-264-534-34
25	83	5.3	657	1	US-08-083-350A-13
26	83	5.3	657	1	US-08-465-500-34
27	83	5.3	657	1	US-08-465-500-34

28	83	5.3	657	2	US-08-346-128-34	Sequence 34, Appl
29	83	5.3	657	3	US-08-532-384-13	Sequence 13, Appl
30	83	5.3	657	3	US-08-893-828-34	Sequence 34, Appl
31	83	5.3	1015	1	US-08-537-210A-1	Sequence 1, Appl
32	83	5.3	1015	1	US-09-113-825-16	Sequence 16, Appl
33	83	5.3	2471	1	US-08-185-432-16	Sequence 16, Appl
34	83	5.3	2471	1	US-08-083-590A-19	Sequence 19, Appl
35	83	5.3	2471	3	US-08-532-384-19	Sequence 19, Appl
36	82	5.2	1093	3	US-08-545-860D-55	Sequence 55, Appl
37	82	5.2	1093	3	PCT-US94-04496-55	Sequence 55, Appl
38	81.5	5.2	371	4	US-09-347-803-27	Sequence 27, Appl
39	81.5	5.2	1026	1	US-08-194-290-7	Sequence 7, Appl
40	80.5	5.1	530	1	US-08-187-793-4	Sequence 4, Appl
41	80	5.1	417	4	US-09-355-115-2	Sequence 2, Appl
42	79.5	5.1	427	4	US-08-886-634-2	Sequence 2, Appl
43	79.5	5.1	427	4	US-09-348-851-2	Sequence 2, Appl
44	79	5.0	425	3	US-08-946-329A-19	Sequence 19, Appl
45	79	5.0	425	4	US-08-567-357A-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-099-676-1
Sequence 1, Application US/09099676
Patent No. 6100075
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/099,676
FILING DATE: HERWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0532 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-855-0572
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 314 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSNO1
CLONE: 2278458
US-09-099-676-1

Query Match 100.0%; Score 1567; DB 3; Length 314;
Best Local Similarity 100.0%; Pred. No. 2.5e-163;
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVGFAGOLAVRFTAGILSAHKIIASSPEMNLPTVSALKRMGVNLTFRSKETVKHSD 60
DB 1 MSVGFAGOLAVRFTAGILSAHKIIASSPEMNLPTVSALKRMGVNLTFRSKETVKHSD 60
QY 61 VLFLAVKPHIIPFLDEIGADVQARHIVVSCAAGVTISSVEKKIMAFQAPAVIRCMNT 120
DB 61 VLFLAVKPHIIPFLDEIGADVQARHIVVSCAAGVTISSVEKKIMAFQAPAVIRCMNT 120
QY 121 PVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGPAYAFMA 180
DB 121 PVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGPAYAFMA 180
QY 181 LDADGVKMGIPRRLAIQLGAQALLGAAKMLDSEQHPQKDNVCSFGATIHAFLE 240
DB 181 LDADGVKMGIPRRLAIQLGAQALLGAAKMLDSEQHPQKDNVCSFGATIHAFLE 240
QY 241 SGGFRSLINAVEASCIRTELQSMADQEKISPAALKKTLDRVKLESPTVSTLTPSSPG 300
DB 241 SGGFRSLINAVEASCIRTELQSMADQEKISPAALKKTLDRVKLESPTVSTLTPSSPG 300
QY 301 KLTLSLALGKKD 314
DB 301 KLTLSLALGKKD 314

RESULT 2

US-09-565-940-1
Sequence 1, Application US/09565910
Patent No. 6288192

GENERAL INFORMATION

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Baughn, Mariah R.

TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA: US/09/565, 910

APPLICATION NUMBER: US/09/565, 910

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/099, 676

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Cerrone, Michael C.

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PF-0532 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-855-0572

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 314 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: PROSNON01
CLONE: 2278458
US-09-565-910-1

Query Match 100.0%; Score 1567; DB 4; Length 314;
Best Local Similarity 100.0%; Pred. No. 2.5e-163;
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVGFAGOLAVRFTAGILSAHKIIASSPEMNLPTVSALKRMGVNLTFRSKETVKHSD 60
DB 1 MSVGFAGOLAVRFTAGILSAHKIIASSPEMNLPTVSALKRMGVNLTFRSKETVKHSD 60
QY 61 VLFLAVKPHIIPFLDEIGADVQARHIVVSCAAGVTISSVEKKIMAFQAPAVIRCMNT 120
DB 61 VLFLAVKPHIIPFLDEIGADVQARHIVVSCAAGVTISSVEKKIMAFQAPAVIRCMNT 120
QY 121 PVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGPAYAFMA 180
DB 121 PVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGPAYAFMA 180
QY 181 LDADGVKMGIPRRLAIQLGAQALLGAAKMLDSEQHPQKDNVCSFGATIHAFLE 240
DB 181 LDADGVKMGIPRRLAIQLGAQALLGAAKMLDSEQHPQKDNVCSFGATIHAFLE 240
QY 241 SGGFRSLINAVEASCIRTELQSMADQEKISPAALKKTLDRVKLESPTVSTLTPSSPG 300
DB 241 SGGFRSLINAVEASCIRTELQSMADQEKISPAALKKTLDRVKLESPTVSTLTPSSPG 300
QY 301 KLTLSLALGKKD 314
DB 301 KLTLSLALGKKD 314

RESULT 3

US-09-099-676-3
Sequence 3, Application US/09099676
Patent No. 6100075

GENERAL INFORMATION

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Baughn, Mariah R.

TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA: US/09/099, 676

APPLICATION NUMBER: US/09/099, 676

FILING DATE: HEREWITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Cerrone, Michael C.

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PF-0532 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-855-0572

TELEX:

RESULT 4
 US-09-565-910-3
 ; Sequence 3, Application US/09565910
 ; Patent No. 6258192
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Baughn, Mariah R.
 ; TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
 ; TITLE OF INVENTION: HOMOLOG
 ; NUMBER OF SEQUENCES: 3
 ;
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ;
 ; COUNTRY: USA
 ; ZIP: 94304
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ;
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FASTSEQ for Windows Version 2.0b
 ;
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/565,910
 ;
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/099,676
 ;
 ; FILING DATE:
 ;
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ceirone, Michael C

RESULT 5
US-08-665-716-2
Sequence 2, Application US/08665716
Patent No. 5789222
GENERAL INFORMATION:
APPLICANT: KELLY, ROSEMARIE
APPLICANT: REGISTER, ELIZABETH A
APPLICANT: MASUREKAR, PRAKASH S
TITLE OF INVENTION: P5C REDUCTASE GENE FROM ZALERION
TITLE OF INVENTION: ARBORICOLA
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: MERCK & CO., INC.
STREET: 126 E. LINCOLN AVENUE
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: US
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665, 716
FILING DATE: 23-JUN-1995

```

? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: KOESEN, ELLIOTT
? REGISTRATION NUMBER: 32,705
? REFERENCE/DOCKET NUMBER: 19453PV
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 908-594-5493
? TELEFAX: 908-594-4720
? INFORMATION FOR SEQ. ID NO.: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 304 amino acids
? type: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
?
US-08-665-716-2

```

Query Match	18.88;	Score 294;	DB 1;	Length 304;
Best Local Similarity	29.98;	Pred. No. 6.2e-24;		
Matches	88;	Conservative 49;	Mismatches 119;	Indels 38;
			Gaps	9

```

QY      1 MASGFIGAOLAVRFPAAGIIISAKHIIISSEPM-----LPT--VSALR-----42
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      11 LTLAVICGCTMGALISGLASLDEIHAPNQSSEIDETPFSKLPYFIACVSPSGAAKI 70
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      43 -----KMGVLTFSNKEVY-KHSDVLELAVPHIIIPILDETG-ADYQAHVIVSCAA 93
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      71 KKALSPKPIPVYKISDNYATACREADVYLTGCKRYMAEGILTEEGEAVDMLKGGKLLISIA 130
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      94 GVLTSSPEKKIAFOAP-----KYVRCKMTNPVVVGEGATVYVATGHAL-VEDGQL 145
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      131 GVPABQVIGYMGKPIPVNPEKEGCLQVVRAMPNTASISRESMTYVITSSPPLSATTSLSI 190
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      146 EQLMSSVGFCTEVEDLIDAVTGLSGSPAFAYMALDA--DGGVYKGLPRRLAIGLGAOA 201
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      191 TWIFKRIQDVVOLPATMTASTALTCGSPAFALITLAAIDGAVAMGLPRAEORMAAQT 250
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY      204 ILGAARKLLDSEQHPCQKNDVCSGPGATIHAFLELSSGFSLLINAYEASCI 257
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      251 MKGAGVIVLSGE-HPALLDKQVTPYGGGCTIGGLVLEEGGVRTGVANARAEATV 303
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

```

RESULT 6
 US-08-605-150A-17
 Sequence 17 Application US/08605150A
 Patent No. 6103520
 GENERAL INFORMATION:
 APPLICANT: Topfer, Reinhard
 APPLICANT: Hausmann, Ludwig
 APPLICANT: Schell, Jozef
 TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE
 TITLE OF INVENTION:
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klein & Szekeres
 STREET: 4199 Campus Drive, Suite 700
 CITY: Irvine
 STATE: CA
 COUNTRY: USA
 ZIP: 92715
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/605,150A
 FILING DATE: 01-MAR-1996
 CLASSIFICATION: 800
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: PCJ\EP94\02936
 FILING DATE: 02-SEP-1994
 PRIORITY APPLICATION DATA:

```

1 APPLICATION NUMBER: DE P4329837.3
2 FILING DATE: 03-SEP-1993
3
4 ATTORNEY/AGENT INFORMATION:
5
6 NAME: Szekeres, Gabor B.
7 REGISTRATION NUMBER: 28,675
8 REFERENCE/DOCKET NUMBER: 542-04-PA
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: 714-854-5502
11 TELEFAX: 714-854-4897
12
13 INFORMATION FOR SEQ ID NO: 17:
14
15 SEQUENCE CHARACTERISTICS
16 LENGTH: 349 amino acids
17 TYPE: amino acid
18 TOPOLOGY: linear
19
20 US-08-605-150A-17

```

Query Match	6.5%;	Score 102;	DB 3;	Length 349;
Best Local Similarity	23.3%;	Pred. No. 0.0083;		
Matches 59;	Conservative 39;	Mismatches 95;	Indels 60;	Gaps 11

[illegible]

RESULT 7
US-08-605-150A-6
Sequence 6, Application US/08605150A
Patent No. 6103520
GENERAL INFORMATION:
APPLICANT: Topfer, Reinhard
APPLICANT: Hausmann, Ludwig
APPLICANT: Schell, Jozef
TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROENASE
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klein & Szekeres
STREET: 4199 Campus Drive, Suite 700
CITY: Irvine
STATE: CA
COUNTRY: USA
ZIP: 92715
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,150A
FILING DATE: 01-MAR-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94\02936
FILING DATE: 02-SEP-1994
PRIOR APPLICATION DATA:


```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,251
; FILING DATE: 14-MAY-1993
; PRIOR APPLICATION DATA: JP 4-147945
; APPLICATION NUMBER: 25,258
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/150/AAOK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 543 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-375-709-17

```

```

Query Match      6.2%; Score 97.5; DB 1; Length 543;
Best Local Similarity 22.5%; Pred. No. 0.053;
Matches 71; Conservative 48; Mismatches 130; Indels 67; Gaps 13;

```

```

QY 8 AGQLAYRFTAGIL-----SAHKTIASSPE-----MNLPTVSALRKMGVNLTRSNK- 53
DB 119 AGILCGSFGAAGLIPSRKALINRIQALPMPGPMENLIHSPSPALERGSVELFLKHV 178
QY 54 ETVKHSDVLEFLAVKPHIIPILDEIGADYQARHIVS-CAAGVTISSVEKKLMAFOPAPK 112
DB 179 RTVEAS--AFGLTPOIYVYRAGLSRDAQGVVGNKVIKASRTVEAKEM--MPAP- 233
QY 113 VIRGNTNPVYVQEGATVATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGS 172
DB 234 -----AKMLQKLVDDGSIYAEQMEALQVPMADDTAEADSG--GH 272
QY 173 GPAVAFMALDADGVKMGIPRRILAIOGAQALIGAAMKLLDSBOHPCOLKDNVCSPGAT 232
DB 273 TDNRPLVTL-----LPTILALKEIQ-----AKYQYDT---PIRVG---CGGGVGT 312
QY 233 IHALHLESGGFRSLLINVAEASCI-----RTRELQSMADQEKISPAALKTKILLDRVYL 286
DB 313 PDAALATFPMGAAYIVTGSINQACVEAGASDHTRKLLATTEMADVMAPAADMEKGVKL 372
QY 287 ESPVTSTLPSSPGKL 302
DB 373 QVVKRGTLFPMRANKL 388

```

RESULT 10

```

; US-08-752-929-17
; Sequence 17, Application US/08752929
; Patent No. 5798259
; GENERAL INFORMATION:
; APPLICANT: YAZAWA, Kazunaga
; APPLICANT: YAMADA, Akiko
; APPLICANT: KATO, Seishi
; APPLICANT: KONDO, Kiyosi
; TITLE OF INVENTION: Gene Coding for Eicosapentaenoic Acid Synthesizing
; TITLE OF INVENTION: Enzymes and Process for Production of Eicosapentaenoic
; TITLE OF INVENTION: Acid
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentlin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,929
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/375,709
; FILING DATE: 20-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,251
; FILING DATE: 14-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-147945
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/150/AAOK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 543 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-752-929-17

```

```

Query Match      6.2%; Score 97.5; DB 1; Length 543;
Best Local Similarity 22.5%; Pred. No. 0.053;
Matches 71; Conservative 48; Mismatches 130; Indels 67; Gaps 13;

```

```

QY 8 AGQLAYRFTAGIL-----SAHKTIASSPE-----MNLPTVSALRKMGVNLTRSNK- 53
DB 119 AGILCGSFGAAGLIPSRKALINRIQALPMPGPMENLIHSPSPALERGSVELFLKHV 178
QY 54 ETVKHSDVLEFLAVKPHIIPILDEIGADYQARHIVS-CAAGVTISSVEKKLMAFOPAPK 112
DB 179 RTVEAS--AFGLTPOIYVYRAGLSRDAQGVVGNKVIKASRTVEAKEM--MPAP- 233
QY 113 VIRGNTNPVYVQEGATVATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGS 172
DB 234 -----AKMLQKLVDDGSIYAEQMEALQVPMADDTAEADSG--GH 272
QY 173 GPAVAFMALDADGVKMGIPRRILAIOGAQALIGAAMKLLDSBOHPCOLKDNVCSPGAT 232
DB 273 TDNRPLVTL-----LPTILALKEIQ-----AKYQYDT---PIRVG---CGGGVGT 312
QY 233 IHALHLESGGFRSLLINVAEASCI-----RTRELQSMADQEKISPAALKTKILLDRVYL 286
DB 313 PDAALATFPMGAAYIVTGSINQACVEAGASDHTRKLLATTEMADVMAPAADMEKGVKL 372
QY 287 ESPVTSTLPSSPGKL 302
DB 373 QVVKRGTLFPMRANKL 388

```

RESULT 11

```

; US-09-090-793-10
; Sequence 10, Application US/09090793
; Patent No. 6140486
; GENERAL INFORMATION:
; APPLICANT: Calgene, LLC
; TITLE OF INVENTION: Production of polyunsaturated fatty acids by expression
; TITLE OF INVENTION: of polyketide-like synthesis genes in plants
; FILE REFERENCE: CGNE.131.0105
; CURRENT APPLICATION NUMBER: US/09/090,793

```

CURRENT FILING DATE: 1998-06-04
 EARLIER APPLICATION NUMBER: 60/048,650
 EARLIER FILING DATE: 1997-06-04
 NUMBER OF SEQ ID NOS: 66
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 10
 LENGTH: 543
 TYPE: PRT
 ORGANISM: Shewanella putrefaciens
 US-09-090-793-10

Query Match 6.2%; Score 97.5; DB 4; Length 543;
 Best Local Similarity 22.5%; Pred. No. 0.053;
 Matches 71; Conservative 48; Mismatches 130; Indels 67; Gaps 13;

QY 8 ACOLAVRTTACIL-----SAHKTIASSPE-----MNLPTYSALKKMGVNLTRSKK- 53
 DB 119 AGILGSGFAGLIPSRVEALINRIQALPNGPYMNLHSPSEPALRGSVELFLKHV 178
 QY 54 ETVKHSDFLAVKPHIIPILDEIGADVQARHIVS--CAAGVTISVEKKIAMFOPAPK 112
 DB 179 RTVES--AFGLTPOIYIYRAAGISRDAGKVVGNKVIAKVSKEVEAKFM--MPAP- 233
 QY 113 VIRCTMTNPVVVQEGATYATGTHALVEDGQLLEQLMSSVGEVEEDLIDAVTGLSGS 172
 DB 234 -----AKMLKLVDDGSTTAEOMELAQVPMADDITAEADSG--GH 272
 QY 173 GAVYAFMALDAGGVKMLPRRLATQLGAQALLGAKKMLDSEHPCOLKDNVCSFGAT 232
 DB 273 TDRPLVTL-----LPTILAKEELQ-----AKYQYDT--PIVVG--CGGSGVT 312
 QY 233 IHALFLESGFRSLINAVEASCI-----RTRELQSMADOEKISPAALKKTLIDRVKL 286
 DB 313 PRAALATFMGAAYIVTGSINGACVACAGASDTRKRLATTEMADVTMAPADMFMGVKL 372
 QY 287 ESPVSTLTTPSSPGKL 302
 DB 373 QVKKRGTLEPMRANKL 388

RESULT 12
 US-08-726-012B-2
 Sequence 2, Application US/08726012B
 Patent No. 5952190
 GENERAL INFORMATION:
 APPLICANT: Hans Joenje, et al.
 TITLE OF INVENTION: CDNA FOR FANCONI ANEMIA COMPLEMENTATION GROUP A
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klarquist Sparkman Campbell Leigh & Whinston, LLP
 STREET: One World Trade Center, Suite 1600, 121 S.W. Salmon Street
 CITY: Portland
 STATE: OR
 COUNTRY: USA
 ZIP: 97204-2988
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Disk, 3.5-inch
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: MS DOS
 SOFTWARE: WordPerfect 5.1+, ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/726,012B
 FILING DATE: 10/04/96
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Richard J. Polley
 REGISTRATION NUMBER: 28,107
 REFERENCE/DOCKET NUMBER: 3812-45520/RJP/DJE
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (503) 226-7391
 TELEFAX: (503) 228-9446
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1455
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-726-012B-2

Query Match 6.1%; Score 95.5; DB 2; Length 1455;
 Best Local Similarity 22.3%; Pred. No. 0.43;
 Matches 79; Conservative 42; Mismatches 130; Indels 103; Gaps 14;

QY 19 GLISAKTIASSPEMNLPYTSALKKMGVNLTRSKKETVKHSDV----- 61
 DB 189 GIYSLQELSHEDMAVGSMLFRNL--CCICEQMEASCOIADVARMLSDPVQMFVLRGP 247
 QY 62 -----FLAVKPHIIP-----FTLDEIGADVQ--ARHIVSCAGVTISSVE 101
 DB 248 OKNSDLRRVEPEKMPQVTVDLQRLMIFALDALAGVDESSTHKIVRCMFVFSGHTL 307
 QY 102 KILMAEQPARK-----VIRCTMTNPV-----VQEGATYATGTHALVED- 141
 DB 308 GSVISTDPLKRFSSHLLTQLITHSPLKASDAVQOMQREMSFARTHPLTSLYRRLVMS 367
 QY 142 -----GOLLQOL-----MSSVGFQ--TEVEEDLIDAVTGLSGSPYAFMALD 182
 DB 368 ABEYLGHLEVELEQEHQMVRLSVFSAIVCPPEAQOLLEDAVARLM-----AQAFESQ 423
 QY 183 ADGVKMGKLPRLATQLGAQALLGAKKMLDS-----EQHPCOLKDNV----- 226
 DB 424 LDSVTAFLVRLQALBGPAPFLSYADMFKASFGSTRGVHGCSSKALVFLFTLSLVEP 483
 QY 227 -SPGATIHALLH-FLSESGFRSLINAVEASCIPTREL-----OSMADOEKISPA 274
 DB 484 ESPRYLOVHILHPLVPSTKRSLLTDTYISLAKTRLADLAKVSTENMGLEYEDLSA 537

RESULT 13
 US-08-605-150A-8
 Sequence 8, Application US/08605150A
 Patent No. 6103520
 GENERAL INFORMATION:
 APPLICANT: Topfer, Reinhard
 APPLICANT: Hausmann, Ludger
 APPLICANT: Schell, Jozeff
 TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klein & Szekeres
 STREET: 4199 Campus Drive, Suite 700
 CITY: Irvine
 STATE: CA
 COUNTRY: USA
 ZIP: 92715
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/605,150A
 FILING DATE: 01-MAR-1996
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/EP94/02936
 FILING DATE: 02-SEP-1994
 PRIOR APPLICATION NUMBER: DE P4329827.3
 FILING DATE: 03-SEP-1993

ATTORNEY/AGENT INFORMATION:
 NAME: Szekeres, Gabor L.
 REGISTRATION NUMBER: 28,675
 REFERENCE/DOCKET NUMBER: 542-04-PA
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 714-854-5502
 TELEFAX: 714-854-4897
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 372 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-605-150A-8

Query Match 6.0%; Score 94.5; DB 3; Length 372;
 Best Local Similarity 23.2%; Pred. No. 0.061; Indels 91; Gaps 18;
 Matches 80; Conservative 52; Mismatches 122;

QY 23 AHKIIASSP-----EMNLPTVSALRKMGVNLTRSN-----52
 DB 37 AAKLIASNTLKLPSFHDEVRMWFEEETLPSEGEKLTIV-INOTNNVNYLPGIKGRNVYA 95
 QY 53 -----KETVKSVDYFLAVKPH-----IIPILDEIGADVQARHIV---VSCAAGVTISS 99
 DB 96 DPDLNNAVKDANML-VFVTPHOFMEGICKRLVGKIQEGAQLSLIKMEVMEGPCMISS 154
 QY 100 VEKKIMAFQAPKVIKCMTPPVVQEGATV-----YATGTHALVDEGQLE---QL 148
 DB 155 LISDLG-----INC-----CYLMGANTANIIEAVEKFESEATVGFRENTDIAEKWVL 201
 QY 149 MSSVGFCTEVEDLIDAVTGLSGSPAVAFMALDAD--GGVKMGLPRLAI-OLGAQALL 205
 DB 202 FSTPYFVMSAVED---VEGVELCGTLKNIVAIAGFVDGLEMMNNTKAIRMIGLEBK 257
 QY 206 GAAKMLDSEQ-----HPCQLKD--NVCSPGATIHAIHFLESQGFSLINAVEASCIR 258
 DB 258 AFSKLLFPYSVKDTPFEESCQVADLITTCIGGRNRKVAEAFKNGGERS--FDLEAEELR 315
 QY 259 TRELSMADQEKISPAALKTKLLDRVKLE-SPTVSTLTPSSPGL 302
 DB 316 GQKLQGVSTAKEV-----YEVLGHRGWLELPFSTVHEISTGRL 355

RESULT 14
 US-08-605-150A-10
 Sequence 10, Application US/08605150A
 Patent No. 6103520
 GENERAL INFORMATION:
 APPLICANT: Topfer, Reinhard
 APPLICANT: Hausmann, Ludwig
 APPLICANT: Schell, Josef
 TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE
 TITLE OF INVENTION: 17
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klein & Szekeres
 STREET: 4199 Campus Drive, Suite 700
 CITY: Irvine
 STATE: CA
 COUNTRY: USA
 ZIP: 92715
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/605,150A
 FILING DATE: 01-MAR-1996
 CLASSIFICATION: 800
 PRIORITY APPLICATION DATA:

APPLICATION NUMBER: PCT/EP94\02936
 FILING DATE: 02-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE P4329827.3
 FILING DATE: 03-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Szekeres, Gabor L.
 REGISTRATION NUMBER: 28,675
 REFERENCE/DOCKET NUMBER: 542-04-PA
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 714-854-5502
 TELEFAX: 714-854-4897
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 372 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-605-150A-10

Query Match 6.0%; Score 94.5; DB 3; Length 372;
 Best Local Similarity 23.2%; Pred. No. 0.061; Indels 91; Gaps 18;
 Matches 80; Conservative 52; Mismatches 122;

QY 23 AHKIIASSP-----EMNLPTVSALRKMGVNLTRSN-----52
 DB 37 AAKLIASNTLKLPSFHDEVRMWFEEETLPSEGEKLTIV-INOTNNVNYLPGIKGRNVYA 95
 QY 53 -----KETVKSVDYFLAVKPH-----IIPILDEIGADVQARHIV---VSCAAGVTISS 99
 DB 96 DPDLNNAVKDANML-VFVTPHOFMEGICKRLVGKIQEGAQLSLIKMEVMEGPCMISS 154
 QY 100 VEKKIMAFQAPKVIKCMTPPVVQEGATV-----YATGTHALVDEGQLE---QL 148
 DB 155 LISDLG-----INC-----CYLMGANTANIIEAVEKFESEATVGFRENTDIAEKWVL 201
 QY 149 MSSVGFCTEVEDLIDAVTGLSGSPAVAFMALDAD--GGVKMGLPRLAI-OLGAQALL 205
 DB 202 FSTPYFVMSAVED---VEGVELCGTLKNIVAIAGFVDGLEMMNNTKAIRMIGLEBK 257
 QY 206 GAAKMLDSEQ-----HPCQLKD--NVCSPGATIHAIHFLESQGFSLINAVEASCIR 258
 DB 258 AFSKLLFPYSVKDTPFEESCQVADLITTCIGGRNRKVAEAFKNGGERS--FDLEAEELR 315
 QY 259 TRELSMADQEKISPAALKTKLLDRVKLE-SPTVSTLTPSSPGL 302
 DB 316 GQKLQGVSTAKEV-----YEVLGHRGWLELPFSTVHEISTGRL 355

RESULT 15
 US-08-614-377A-7
 Sequence 7, Application US/08614377A
 Patent No. 5976864
 GENERAL INFORMATION:
 APPLICANT: Smit, John
 APPLICANT: Bingle, Wade H.
 APPLICANT: No. 5976864ellini, John F.
 TITLE OF INVENTION: EXPRESSION AND SECRETION OF
 TITLE OF INVENTION: HETEROLOGOUS
 TITLE OF INVENTION: POLYPEPTIDES FROM CAULOBACTER
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson PC
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PC-DOS/MS-DOS

```

: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/614,377A
: FILING DATE: 12-MAR-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/194,290
: FILING DATE: 09-FEB-1994
: CLASSIFICATION: 435
: APPLICATION NUMBER: US 07/895,367
: FILING DATE: 09-JUNE-1992
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: TS&O, Y. ROCKY
: REGISTRATION NUMBER: 34053
: REFERENCE/DOCKET NUMBER: 08106/002001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-542-5070
: TELEFAX: 617-542-8906
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1026 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-614-377A-7

```

```

Query Match      5.8%; Score 91.5; DB 2; Length 1026;
Best Local Similarity 22.7%; Pred. No. 0.67;
Matches 57; Conservative 37; Mismatches 104; Indels 53; Gaps 10;

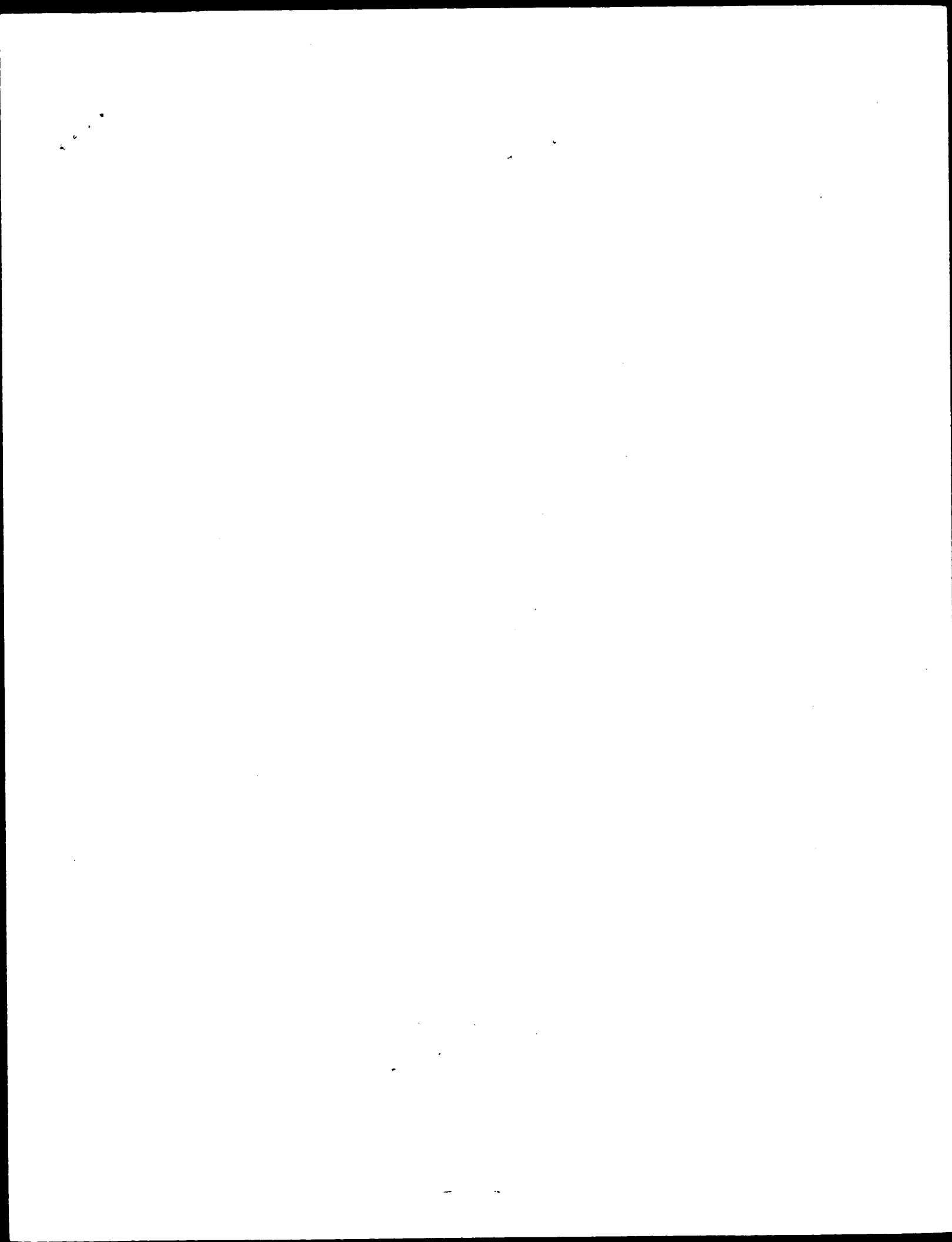
```

```

QY 27 IASSPEMNLPTVSALKRMGVNLTNRNKTETKHSVDVLEAVKPHIIFILDELIGADVOARH 86
    :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
DB 768 LSSSALAAGTVALAGVETVNIATDTNTTAHVDTLQA-----TSAKS 812
    :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 87 IVVSCAAGVTISSV-EKKLMAFQ-----PARKVIRKMTNP---VVVQEGATVYATGT 135
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
DB 813 IVVTGAGLNLNTGNTAVTSFDASAVTGTAPAVTFVSANTTVGEVYVIRGA-----GA 867
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
QY 136 HALVEPQQLLEQIMSVGFCTEVEEDLIDAVTGLSGS-----GPAVAFMAL-DADGG 186
    :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
DB 868 DSLTGSATANDTIIGGAGADTLVYGGTDTFTGCGADIEDINAIGTSTAFYTTIDAAG 927
    :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 187 VKM---GLPRRLAI---QLGAQALLGAAMKMLDSEQHPCQLKDNVCSPGATTHALHFL 240
    :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
DB 928 DKLDLVGISTNGAIDGAFGAAVTLGAATL-----AOYLDAAGADGSGSTVAKKWFQ 980
    :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 241 SGGFRSLTINA 251
    |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
DB 981 FGGDTIYVYVDS 991

```

Search completed: January 31, 2002, 14:53:52
Job time: 22 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:54:15 ; Search time 12.98 Seconds

(without alignments)
886.961 Million cell updates/sec

Title: US-09-912-717-1

Perfect score: 1567
Sequence: 1 MSVFGTGAQLAYRTAGI.....TPSSPGKLLTRSLALGKKD 314

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1306	83.3	319	1	PROG_HUMAN
2	512.5	32.7	276	1	PROG_ARATH
3	501.5	32.0	274	1	PROG_SOYBN
4	479	30.6	273	1	PROG_PEA
5	476	30.4	278	1	PROG_ACTCH
6	421	26.9	267	1	PROG_SYNY3
7	402.5	25.7	269	1	PROG_ECOLI
8	377	24.1	272	1	PROG_PSEAE
9	370	23.6	270	1	PROG_CORGL
10	345	22.0	270	1	PROG_BACSL
11	326.5	20.8	294	1	PROG_MYCLE
12	324.5	20.7	278	1	PROG_BACSL
13	323.5	20.6	261	1	PROG_THERM
14	306.5	19.6	295	1	PROG_MYCTU
15	301.5	19.2	311	1	PROG_NEUCR
16	297	19.0	265	1	PROG_AQUAE
17	296	18.9	320	1	PROG_ZALAR
18	281.5	18.0	286	1	PROG_YEAST
19	279	17.8	293	1	PROG_TREPA
20	266	17.0	299	1	PROG_CABEL
21	264	16.8	271	1	PROG_HAETN
22	252.5	16.1	278	1	PROG_VIBAL
23	193.5	12.3	257	1	PROG_HELPY
24	179.5	11.5	257	1	PROG_HELPY
25	163.5	10.4	273	1	CME4_BACSL
26	109	7.0	599	1	CENB_MOUSE
27	106	6.8	348	1	GPDA_RABIT
28	104	6.6	272	1	YKFA_BACSL
29	104	6.6	451	1	YVBA_METTF
30	102	6.5	251	1	PROG_METSM
31	102	6.5	348	1	GPDA_MOUSE
32	101	6.4	599	1	CENB_HUMAN
33	100	6.4	348	1	GPDA_RAT

34	97	6.2	272	1	THIM_BACSL
35	97	6.2	348	1	GPDA_HUMAN
36	97	6.2	351	1	GPDA_FUGRU
37	95.5	6.1	1455	1	FACA_HUMAN
38	91.5	5.8	346	1	ARCC_BACSL
39	91.5	5.8	347	1	IDIT2_ERKHE
40	91	5.8	790	1	SYFB_CHIMU
41	90	5.7	397	1	TRB1_AQUAE
42	89.5	5.7	371	1	MAK_ECOLI
43	89.5	5.7	569	1	AMP2_LYCES
44	89.5	5.7	2541	1	TALI_HUMAN
45	89	5.7	906	1	KBP2_CHICK

ALIGNMENTS

RESULT	ID	PROG_HUMAN	STANDARD	PRT	319 AA.
AC	P32322				
DT	01-OCT-1993	(Rel. 27, Created)			
DT	01-OCT-1993	(Rel. 27, Last sequence update)			
DT	20-AUG-2001	(Rel. 40, Last annotation update)			
DE	PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).				
GN	PYCRL				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92112821; Pubmed=1730675;				
RA	Dougherty K.M., Brandtiss M.C., Valle D.;				
RT	*Cloning human pyrrolidine-5-carboxylate reductase cDNA by				
RT	complementation in Saccharomyces cerevisiae.*				
RL	J. Biol. Chem. 267:871-875(1992).				
CC	-1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = L-PYRROLINE-5-CARBOXYLATE + NAD(P)H.				
CC	-1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.				
CC	-1- SUBUNIT: HOMODECAMER OR HOMODODECAMER.				
CC	-1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: M77836; AAA36407.1; -.				
DR	PIR: A41770; A41770.				
DR	MIM: 179035; -.				
DR	InterPro: IPR000304; P5CR.				
DR	Pfam: PF01089; P5CR.1.				
DR	PROSITE: PS00521; P5CR.1.				
KW	Oxidoreductase; proline biosynthesis; NADP.				
SQ	SEQUENCE 319 AA; 33374 MW; F5E74B5BDF475EF CRC64;				

Query Match 83.3%; Score 1306; DB 1; Length 319;
Best Local Similarity 82.7%; Pred. No. 1.5e-89;
Matches 263; Conservative 27; Mismatches 22; Indels 6; Gaps 2;

QY	1	MSVFGTGAQLAYR---FTAGILSAHKITASSPENLPTVSALRKMGVLTNRNKETV 56
DB	1	MSVFGTGAQLAYR---FTAGILSAHKITASSPENLPTVSALRKMGVLTNRNKETV 56
QY	57	KHSVFLAVKPHIPIPIIDIGADVAGARIVVSCAGVYITSSVEKKLMAFORAPVIRC 116
DB	61	QHSVFLAVKPHIPIPIIDIGADVAGARIVVSCAGVYITSSVEKKLMAFORAPVIRC 120

QY 117 MNTPEVVOGADPVATGTHALVEEDGOLLEOLMSSVGFTEVEEDLIDAVTGLSGSPAY 176
 DB 121 MNTPEVVOGADPVATGTHALVEEDGOLLEOLMSSVGFTEVEEDLIDAVTGLSGSPAY 180
 QY 177 AFMALD--ADGGVKGMLPRLAIGLGAQALLGAARKMLDSEOHPCOLKDNVCSPGATIH 234
 DB 181 AFTALDALADGKMGMLPRLAIGLGAQALLGAARKMLDSEOHPCOLKDNVCSPGATIH 240
 QY 235 ALHFLSEGGFRSLINAVASCIPTRELOSMDQEKISPAALKTLLDRKLESTVSTL 294
 DB 241 ALHFLSEGGFRSLINAVASCIPTRELOSMDQEKISPAALKTLLDRKLESTVSTL 300
 QY 295 TPSSPGKLTLSLALGK 312
 DB 301 SPGGHKLPLRSLAPAGK 318

RESULT 2
 ID PROC ARATH STANDARD: PRT: 276 AA.
 AC P54904:
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
 GN PROCI OR AT5G14800 OR T9L3.100.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94294559; PubMed=8022935;
 RA Verduygen N., Villarroel R., van Montagu M.;
 RT "Osmoregulation of a pyrroline-5-carboxylate reductase gene in
 RL Arabidopsis thaliana.";
 RL Plant Physiol. 103:771-781(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. LANDSBERG ERRECTA;
 RA Verduygen N., Villarroel R., Hua X., van Montagu M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-21016721; PubMed=11130714;
 RX Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naro K., Okumura S., Shino S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Milvanev E., Ozeresky P., Riley A., Strommatt C.,
 RA Wagner-McPherson C., Wolam A., Yeakum M., Bell M., Dedina J.,
 RA Parnell L., Shah R., Rodriguez M., Hoan See L., Vill D., Baker J.,
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Marienissen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
 RA Entlan K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Gijmenez B., Zimmermann W.,
 RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirke W., Mooljman P., Klein Lankhorst R.,
 RA Weitzenecker T., Bothe G., Rose M., Hauf J., Bernerstorfer S., Hempel S.,
 RA Feilgenhans G., Lamberth S., Villarroel R., Gietlen J., Arlides W.,
 RA Behts O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 RA Scheller C., Zaccaria P., Mewes H.-W., Beyer M., Fritsch P.F.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis

thaliana.";
 RL Nature 408:823-826(2000).
 CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = L-PYRROLINE-5-
 CC CARBOXYLATE + NAD(P)H.
 CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M76538; AAA61346.1; -;
 CC EMBL: Y08951; CAA70148.1; -;
 CC EMBL: AL391149; CAC01879.1; -;
 CC HSP: P80028; 170F.
 CC InterPro: IPR000304; P5CR.
 CC Pfam: PF01089; P5CR.1.
 CC PROSITE: PS00521; P5CR.1.
 CC Oxidoreductase; Proline biosynthesis; NADP.
 KW SEQUENCE 276 AA; 28624 MW; B577A01C92A3A28B CRC64;
 SQ

Query Match 32.7%; Score 512.5; DB 1; Length 276;
 Best Local Similarity 42.5%; Pred. No. 4,8e-31;
 Matches 113; Conservative 52; Mismatches 90; Indels 11; Gaps 4;

QY 3 VGFAGAGIAYRFR---TAAGIISAHKTIASPEPMNPTVSALRKMGNVTRSNKETVKH 58
 DB 13 VGFAGAGIAYRFR---TAAGIISAHKTIASPEPMNPTVSALRKMGNVTRSNKETVKH 71

QY 59 SDVLFPAVKPHIPIFIDEGADYQARHIVSCAGVYISVEKKLMAFQAPADRVRCMT 118
 DB 72 SDVVFISVKKPVYKAVTELKSKLKNKILVSAAGIKLNDLDE---WSGDQRFIVMP 127

QY 119 NTPVVOGADPVATGTHALVEEDGOLLEOLMSSVGFTEVEEDLIDAVTGLSGSPAYAV 178
 DB 128 NTPVVOGADPVATGTHALVEEDGOLLEOLMSSVGFTEVEEDLIDAVTGLSGSPAYAV 187

QY 179 MALD--ADGGVKGMLPRLAIGLGAQALLGAARKMLDSEOHPCOLKDNVCSPGATIH 236
 DB 188 LALDALADGKMGMLPRLAIGLGAQALLGAARKMLDSEOHPCOLKDNVCSPGATIH 247

QY 237 HFLSEGGFRSLINAVASCIPTRELOSMDQEKISPAALKTLLDRKLESTVSTL 294
 DB 248 HFLSEGGFRSLINAVASCIPTRELOSMDQEKISPAALKTLLDRKLESTVSTL 300

RESULT 3
 ID PROC SOYBN STANDARD: PRT: 274 AA.
 AC P17817;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
 OS Glycine max (Soybean).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90340278; PubMed=2199815;
 RA Delaney A.J., Verma D.P.S.;
 RT "A soybean gene encoding delta 1-pyrroline-5-carboxylate reductase
 RT was isolated by functional complementation in Escherichia coli and is
 RT found to be osmoregulated.";

RU Mol. Genet. 221:299-305(1990).
CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = 1-PYRROLINE-5-CARBOXYLATE + NAD(P)H.
CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES, BUT MOSTLY IN NODULES.
CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X16352; CAA34401.1; -.
CC PIR: S10186; S10186.
CC InterPro: IPR000304; P5CR.
CC Pfam: PF01089; P5CR; 1.
CC PROSITE: PS00521; P5CR; 1.
CC Oxidoreductase; Proline biosynthesis; NADP.
CC SEQUENCE 274 AA; 28586 MW; 933CFCFD07598B63 CRC64;
KW

Query Match 32.0%; Score 501.5; DB 1; Length 274;
Best Local Similarity 41.6%; Pred. No. 3.1e-30;
Matches 111; Conservative 51; Mismatches 94; Indels 11; Gaps 4;

QY 2 SVGTGAGQLAYR---TAAGILSAHKIIASSPEMLPTVSALKRMGVNLTNRKKEV 57
DB 12 TLGTGAGKMAESIRKAGVRSGLVPPSR-IRTAHVHNLARKGAFESFVTVLPSSNDVVR 70
DB 58 HSDVFLAVKPHIIFILDEIGADVQARHIYSCAGVTISSVEKKLMAFPAPAKVIR 117
DB 71 ESDVAVLVYKPOLVDVYVSKLPLLTJKHLLVSAVAGTKL---KDLQEMGNDEFIRVM 126
QY 118 TPTPVVOGAVVYVGTHTALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSPPA 177
DB 127 PTPPAVVOGAVVYVGTHTALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSPPA 186
QY 178 FMALD-ADGVYKMGIPRLAIIQIGAQALLGAAKMLDSEQHPCOLKDNVSPGATTHA 235
DB 187 YLAIEALADGVAAAGLPRLSLSLASQVTLGAASVSGTKHPCOLKDNVSPGATTHG 246
QY 236 LHFLESIGFRSLIINAVASCIPTREL 262
DB 247 IHELENGGFRGTLMAVVAARSRREL 273
RESULT 4
PROC_PEA STANDARD: PRT: 273 AA.
AC 004708;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
GN PROC.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RA Williamson C.L., Slocum R.D.;
RT "Molecular cloning and evidence for osmoregulation of the P1-
RT pyrroline-5-carboxylate reductase (proc) gene in pea (Pisum sativum
RT L.)";
RL Plant Physiol. 100:1464-1470(1992).

CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = 1-PYRROLINE-5-CARBOXYLATE + NAD(P)H.
CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X62842; CAA44646.1; -.
CC InterPro: IPR000304; P5CR.
CC Pfam: PF01089; P5CR; 1.
CC PROSITE: PS00521; P5CR; 1.
CC Oxidoreductase; Proline biosynthesis; NADP.
CC SEQUENCE 273 AA; 28237 MW; D719F52935966588 CRC64;
KW

Query Match 30.6%; Score 479; DB 1; Length 273;
Best Local Similarity 40.1%; Pred. No. 1.4e-28;
Matches 108; Conservative 55; Mismatches 90; Indels 16; Gaps 5;

QY 2 SVGTGAGQLAYR---TAAGILSAHKIIA--SPEMLPTVSALKRMGVNLTNRKKEV 55
DB 12 TLGTGAGKMAESIRKAGVRSGLVPPSR-IRTAHVHNLARKGAFESFVTVLPSSNDVVR 67
QY 56 VKHSDVFLAVKPHIIFILDEIGADVQARHIYSCAGVTISSVEKKLMAFPAPAKVIR 115
DB 68 VRASNVVYVSKPOLVDVYVSKLPLLTJKHLLVSAVAGTKL---KDLQEMGNDEFIRVM 123
QY 116 TPTPVVOGAVVYVGTHTALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSPPA 175
DB 124 VPTTPPAVVOGAVVYVGTHTALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSPPA 183
QY 176 YAFMALD-ADGVYKMGIPRLAIIQIGAQALLGAAKMLDSEQHPCOLKDNVSPGATTH 233
DB 184 YLAIEALADGVAAAGLPRLSLSLASQVTLGAASVSGTKHPCOLKDNVSPGATTH 243
QY 234 HALFLESIGFRSLIINAVASCIPTREL 262
DB 244 AGVHELENGGFRGTLMAVVAARSRREL 272

RESULT 5
PROC_ACTCH STANDARD: PRT: 278 AA.
AC 004016;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
GN PROC.
OS Actinidia chinensis (kiwifruit) (Yangtze).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asceridae; Ericales; Actinidiaceae; Actinidia.
OX NCBI_TaxID=3625;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Axillary bud;
RC Walton E.F., Podivinsky E., Wu R.M., Reynolds P.H.S., Young L.W.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = 1-PYRROLINE-5-CARBOXYLATE + NAD(P)H.
CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE FAMILY.
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration

RP SEQUENCE FROM N.A.
BA Duncan M : Allen F : Araujo B : Baricchio A M : Chung F : Davis K

Db 237 EAAKSFQANGFEALVEQALNNAQSRAELAQLOQ 272

RESULT 9

PROC_CONGL STANDARD; PRT; 270 AA.

ID PROC_CONGL

AC P46540;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).

OS PROC.

OS Corynebacterium glutamicum (Brevibacterium flavum).

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;

OC Actinomycetales; Corynebacteriaceae;

OC Corynebacterium.

OC NCBI_TaxID=1718;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 17965 / Melassecola;

RA MEDLINE=96345604; PubMed=8755867;

RA Anki S., Serebrijski I., Reyes O., Leblon G.;

RT "Mutations in the Corynebacterium glutamicum proline biosynthetic

RT pathway: a natural bypass of the pro step.";

RL J. Bacteriol. 178:4412-4419(1996).

CC -1 CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = L-PYRROLINE-5-

CC CARBOXYLATE + NAD(P)H.

CC -1 PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.

CC -1 SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE

CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC DR EMBL; U31225; AAC4172.1; -

CC DR InterPro; IPR000304; P5CR.

CC DR Pfam; PF01089; P5CR; 1.

CC DR PROSITE; PS00521; P5CR; 1.

CC KW Oxidoreductase; Proline biosynthesis; NADP.

CC SEQUENCE 270 AA; 28223 MW; 90150E233D94158F CRC64;

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

PROH_BACSU STANDARD; PRT; 270 AA.

AC P14363; 031828;

DT 01-JAN-1990 (Rel. 13, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE PYRROLINE-5-CARBOXYLATE REDUCTASE HOMOLOG 1.

GN YOKE OR PROC.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OC NCBI_TaxID=1423;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;

RT Submitted (NOV-1997) to the EMBL/GenBank/DBD databases.

RL [2]

RP SEQUENCE OF 1-256 FROM N.A.

RC STRAIN=W23; AND 168;

RA MEDLINE=91192601; PubMed=1849493;

RA Ahn K.S., Wake R.G.;

RT "Variations and coding features of the sequence spanning the

RT replication terminus of Bacillus subtilis 168 and W23 chromosomes.";

RL Gene 98:107-112(1991).

CC [3]

CC SEQUENCE OF 1-200 FROM N.A.

CC STRAIN=W23;

RA MEDLINE=89155440; PubMed=2493444;

RA Lewis P.J., Wake R.G.;

RT "DNA and protein sequence conservation at the replication terminus in

RT Bacillus subtilis 168 and W23.";

RL J. Bacteriol. 171:1402-1408(1989).

CC -1 FUNCTION: NOT KNOWN, BUT COULD BE INVOLVED IN GLUTAMINE

CC BIOSYNTHESIS.

CC -1 SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE

CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC DR EMBL; Z99114; CAB13741.1; -

CC DR EMBL; M24523; AAA22722.1; -

CC DR PIR; B32807; B32807.

CC DR Subtilist; Bg11049; YOKE.

CC DR InterPro; IPR000304; P5CR.

CC DR Pfam; PF01089; P5CR; 1.

CC DR PROSITE; PS00521; P5CR; 1.

CC KW Hypothetical protein; Oxidoreductase; Complete proteome.

CC SEQUENCE 270 AA; 29402 MW; C7B9E1680BFF9A7 CRC64;

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

PROH_BACSU STANDARD; PRT; 270 AA.

AC P14363; 031828;

DT 01-JAN-1990 (Rel. 13, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE PYRROLINE-5-CARBOXYLATE REDUCTASE HOMOLOG 1.

GN YOKE OR PROC.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OC NCBI_TaxID=1423;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;

RT Submitted (NOV-1997) to the EMBL/GenBank/DBD databases.

RL [2]

RP SEQUENCE OF 1-256 FROM N.A.

RC STRAIN=W23; AND 168;

RA MEDLINE=91192601; PubMed=1849493;

RA Ahn K.S., Wake R.G.;

RT "Variations and coding features of the sequence spanning the

RT replication terminus of Bacillus subtilis 168 and W23 chromosomes.";

RL Gene 98:107-112(1991).

CC [3]

CC SEQUENCE OF 1-200 FROM N.A.

CC STRAIN=W23;

RA MEDLINE=89155440; PubMed=2493444;

RA Lewis P.J., Wake R.G.;

RT "DNA and protein sequence conservation at the replication terminus in

RT Bacillus subtilis 168 and W23.";

RL J. Bacteriol. 171:1402-1408(1989).

CC -1 FUNCTION: NOT KNOWN, BUT COULD BE INVOLVED IN GLUTAMINE

CC BIOSYNTHESIS.

CC -1 SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE

CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC DR EMBL; Z99114; CAB13741.1; -

CC DR EMBL; M24523; AAA22722.1; -

CC DR PIR; B32807; B32807.

CC DR Subtilist; Bg11049; YOKE.

CC DR InterPro; IPR000304; P5CR.

CC DR Pfam; PF01089; P5CR; 1.

CC DR PROSITE; PS00521; P5CR; 1.

CC KW Hypothetical protein; Oxidoreductase; Complete proteome.

CC SEQUENCE 270 AA; 29402 MW; C7B9E1680BFF9A7 CRC64;

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

```

Query Match          20.8%; Score 326.5; DB 1; Length 294;
Best Local Similarity 36.2%; Pred. No. 2,8e-17;
Matches 88; Conservative 45; Mismatches 77; Indels 33; Gaps 7

QY 48 LTRSNKETVHSHDYLAVLAKPHIIPILDEI-----GADYQARHIVSCAAGVITISYER 102
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 54 LVTSVTADEAMVAVVAVKPTDVE5VMDLVOAAAVANDSAEGVLTAVAGVITYLE 113
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 103 KLMEFQAPVPIRCMTPTPVYVQEGATVATGTHALVEGCLLEBMS---SVGFCTEVE 159
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 114 KLPGTPT---VVRMPNAAALVAGAVTIAKRGV---TGOQFEDVLAMPDAVGVLTVP 167
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 160 EDLIALVTGLSGSPAPAFEMALDA--DGGVKKGLPRRLATQLGADALGAACKML---DS 214
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 168 ESQADAVTAVSGSPAFVFLVEALVDAGVAVGLTRQVATETLAQTMGSAAMLEMDQ 227
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 215 EQH-----PCQLADNCSPEGATIIHALHFLPESGFRSLILINVEASCIPTRT 260
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 228 DRHSAEVAPLGAVDPAQLRATITIPSGTTPAALRELERGRLMRVDAVVAQAARKIRSE 287
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 261 ELQ 263
      | :
DB 288 QLR 290

RESULT 12
PROL_BACSU          STANDARD;          PRT;          278 AA.
ID PROL_BACSU
AC P54552;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PYRROLINE-5-CARBOXYLATE REDUCTASE HOMOLOG 2.
GN YQJO.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
CX NCBI_TaxID=1423;
KM
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RA Kobayashi Y., Mizuno M., Masuda S., Takemaru K., Hosono S.,
   Sato T., Takeuchi M.
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
   FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
```


DR Pfam: PF01089; P5CR: 1.
 DR PROSITE: PS00521; P5CR: 1.
 KW Oxidoreductase; Proline biosynthesis; NADP: Complete proteome.
 SQ SEQUENCE 295 AA; 30171 MW; F95C3f407BE5408F CXC64;

Query Match 19.6%; Score 306.5; DB 1; Length 295;
 Best Local Similarity 34.3%; Pred. No. 8.5e-16;
 Matches 86; Conservative 39; Mismatches 91; Indels 35; Gaps 6;

QY 48 LTRSNKETVKSVDLFLAVKP-HIIPFIDEIGADV-----QARIIVSCAGATISSVE 101
 DB 54 LVTASADVEMNTFVVAVKPADVEPIADLANAANAENDSACQVEFVVAAGITIAVFE 113
 QY 102 KTLMAFQAPKPYRCMTNTPVVQSGATVYATGTHALVEDQQLLEQLMSVGFCTEVED 161
 DB 114 SKLPAGTP---VVRAMPVMAALVGCAGYALAKGREVTPQGLEEVSALEFDVAGVLTYPES 170
 QY 162 LIDAVTGLSGSPAYAFMALD--DGGVKGILPRRLATQLOALGAAKMLDSEQH-- 217
 DB 171 QLDATVAVSGSPATFLLVEALVDAGVGLSRQVATDIAAQTMAASAMLERMEQDQ 230
 QY 218 -----PCQLKDNVCSPGCATIHALFLESGGFRSLINAVEASCIPTREL 262
 DB 231 GGANGELMGLRYDLTASHLRAAVTSPGGTTAAALRELGRGFRMAVDAVAQAA----- 283
 QY 263 GSMADOEKISP 273
 DB 264 KRSSEQLRITP 294

RESULT 15
 PROC_NEUCR ID PROC_NEUCR STANDARD; PRT: 311 AA.

AC Q12641;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
 GN PRO-1.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariiales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96004701; PubMed=7565596;
 RA Davis C.R., McPeak M.A., McClung C.R.;
 RT "Molecular characterization of the proline-1 (pro-1) locus of
 RT Neurospora crassa, which encodes delta 1-pyrroline-5-carboxylate
 RT reductase."
 RL Mol. Gen. Genet. 248:341-350(1995).
 CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = L-PYRROLINE-5-
 CC CARBOXYLATE + NAD(P)H.
 CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U30317; AAA83568.1;
 CC InterPro: IPR000304; P5CR.
 DR Pfam: PF01089; P5CR: 1.
 DR PROSITE: PS00521; P5CR: 1.
 KW Oxidoreductase; Proline biosynthesis; NADP.
 SQ SEQUENCE 311 AA; 32151 MW; 9A6D525D845984DF CRC64;

Query Match 19.2%; Score 301.5; DB 1; Length 311;
 Best Local Similarity 31.8%; Pred. No. 2.1e-15;
 Matches 81; Conservative 49; Mismatches 102; Indels 23; Gaps 6;

QY 22 SAHKTIASSPEENKLPVLSALRRKGNLIRSNKETVKSVDLFLAVKPHIIPFIDEIG-- 79
 DB 43 SVAKVESALSPLVKRSVSTLR---VLQSTSNVSAANAHEADILLQCKPVMVSGLLSAGMK 99
 QY 80 -----ADVQARRHYVSCAGATISSVEKL-----MAFQAPKPYRCMTNTPVVQ 125
 DB 100 DALTYKHTGTHARSKRIIISICAGVTPDLERVLREDVGLSADNLPVVRAMPNTASKIR 159
 QY 126 ECATYVATGTHALVED--GQLLEQLMSVGFCTEVEEDLIDAVTGLSGSPAYAFMALD-- 182
 DB 160 ESMPTVINTVDPPLPDTVTLLTWIFERIGEVVYLPPIHMDACTSLCASGTAFAFALMEAA 219
 QY 183 ADGGVKGILPRRLATQLOALGAAKMLDSEQHPCQLKDNVCSPGCATIHALFLESG 242
 DB 220 ADGGVAMGLPRAENRMAAQTMRGAAGLVEGE-HPAILREKVSTPGCTIGGLVLEEG 278
 QY 243 GFRSLINAVEASCI 257
 DB 279 GYRAAVARAVEREATV 293

Search completed: January 31, 2002, 14:57:38
 Job time: 203 sec

Thu Jan 31 15:07:58 2002

us-09-912-717-1.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:53:30 ; Search time 16.15 Seconds

(without alignments)
1481.041 Million cell updates/sec

Title: US-09-912-717-1

Perfect score: 1567
Sequence: 1 MSVGFAGQLAYRFTAGT.....TPSSPGKLLTRSLALGKKD 314

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : PIR_68:*

1: PIR1:*\n2: PIR2:*\n3: PIR3:*\n4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1306	83.3	319	1 A41770	pyrroline-5-carbox
2	578	36.9	279	2 T23765	hypothetical prote
3	512.5	32.7	276	1 J02334	pyrroline-5-carbox
4	501.5	32.0	274	1 S10186	pyrroline-5-carbox
5	479	30.6	273	2 T06477	probable pyrroline
6	421	27.0	284	2 T36286	pyrroline-5-carbox
7	406.5	25.9	267	2 S76767	probable pyrroline
8	402.5	25.7	269	1 H85533	pyrroline-5-carbox
9	377	24.1	273	2 J00418	pyrroline-5-carbox
10	345	22.0	270	2 E68682	pyrroline-5-carbox
11	326.5	20.8	294	2 S72897	pyrroline-5-carbox
12	324.5	20.7	278	1 G69964	pyrroline-5-carbox
13	323.5	20.6	261	1 JC2078	pyrroline-5-carbox
14	323.5	20.6	279	2 G83837	pyrroline-5-carbox
15	319	20.4	255	2 E72360	pyrroline-5-carbox
16	306.5	19.6	295	2 G70745	probable proc prot
17	301.5	19.2	311	2 S57863	pyrroline-5-carbox
18	297.5	19.0	305	2 C82524	pyrroline-5-carbox
19	297.5	18.9	265	2 F70315	pyrroline carboxyl
20	296	18.9	320	2 JC4830	pyrroline-5-carbox
21	296	18.9	262	2 D86860	pyrroline-5-carbox
22	295.5	18.7	266	2 G72769	Delta 1-pyrroline-
23	293	18.7	282	2 T50305	pyrroline-5-carbox
24	292.5	18.4	264	2 C75385	pyrroline-5-carbox
25	289	18.0	266	2 S25293	probable pyrroline
26	281.5	17.8	263	2 D71281	hypothetical prote
27	279	17.0	299	2 T29226	pyrroline-5-carbox
28	266	16.8	271	2 I64060	
29	264				

30	258.5	16.5	272	2 D82321	pyrroline-5-carbox
31	229	14.6	267	2 G83760	pyrroline-5-carbox
32	212.5	13.6	263	2 H82015	pyrroline pyrroline
33	212.5	13.6	263	2 E81243	pyrroline-5-carbox
34	193.5	12.3	257	2 F64664	pyrroline-5-carbox
35	192.5	12.3	243	2 G81310	probable pyrroline
36	179.5	11.5	257	2 E71850	probable pyrroline
37	179	11.4	275	2 C83816	non-essential gene
38	162.5	10.4	273	1 F69602	pyrroline-5-carbox
39	148.5	9.5	253	2 H69219	pyrroline-5-carbox
40	114.5	7.3	264	1 F69100	conserved hypotet
41	106.5	6.8	1043	2 H83329	probable RND efflu
42	106	6.8	348	2 A32512	glycerol-1-3-phospha
43	104	6.6	272	1 F69855	pyrroline-5-carbox
44	104	6.6	451	2 S26441	hypothetical prote
45	102	6.5	349	2 A25952	glycerol-1-3-phospha

ALIGNMENTS

RESULT 1

A41770
pyrroline-5-carboxylate reductase (EC 1.5.1.2) - human

N:Alternate names: P5C reductase

C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A41770

R:Dougherty, K.M.; Brandtiss, M.C.; Valle, D.

J. Biol. Chem. 267, 871-875, 1992.

A:Title: Cloning human pyrroline-5-carboxylate reductase cDNA by complementation in S

A:Reference number: A41770; MUID:92112821

A:Accession: A41770

A:Molecule type: mRNA

A:Residues: 1-319 <DOU>

A:Cross-references: GB:M77836; NID:q189497; PIDN:AAA36407.1; PID:q189498

A>Note: sequence extracted from NCBI Backbone (NCBIN:75606, NCBI:P75608)

A:Gene: GDB:PYCRL; P5C; PYCR

A:Cross-references: GDB:135716; OMIM:179035

A:Map position: 17pter-17qter

C:Superfamily: pyrroline-5-carboxylate reductase

C:Keywords: oxidoreductase; proline biosynthesis

Query Match 83.3%; Score 1306; DB 1; Length 319;
Best Local Similarity 82.7%; Pred. No. 2.1e-96;
Matches 263; Conservative 27; Mismatches 22; Indels 6; Gaps 2;

QY	1	MSVGFAGQLAYR---FTAGILSAHKITASSPENMLPTVSALRKMGVNLFRSNKETV	56
DB	1	MSVGFAGQLAFALAKFETAGVLAHKIMASSPMDLAVSALRKMGVNLFRHNKETV	60
QY	57	KHSVDYFLAVRPHIIFLDEIGADVQARHIVSCAGVTTSSVEKKLMAFOPAPKPIRC	116
DB	1	QHSVDYFLAVRPHIIFLDEIGADVQARHIVSCAGVTTSSVEKKLMAFOPAPKPIRC	120
QY	117	MTNPPVVOEGATVYATGTHALVEDGOLLEOLMSSVGCFTVEEDLLIDAVTGLSGSPAY	176
DB	121	MTNPPVVRREGATVYATGTHAIVEDGRLMEQLLSTVGCTEVEEDLLIDAVTGLSGSPAY	180
QY	177	AFMALD--ADGVNKGFLPRRLAIOGAQALGAAMKMLDSBOHRCQKLDVNCSPGATIH	234
DB	181	AFTALDALADGVNKGFLPRRLAIVRGAQALGAAMKMLHDSBOHRCQKLDVNCSPGATIH	240
QY	235	ALHPLFESGFRSLINAVEASCIRTRLOSMADDEKISPAALKTLLDRKLESPYSTL	294
DB	241	ALHPLFESGFRSLINAVEASCIRTRLOSMADDEKISPAALKTLLDRKLESPYSTL	300
QY	295	TPSSPGKLLTRSLALGK 312	
DB	301	SPSGHKKLLPRSLAPACK 318	

Db 247 IHELENGCFRGITLMNVAAAKRSREL 273

RESULT 5

probable pyrroline-5-carboxylate reductase (EC 1.5.1.2) - garden pea

C:Species: Pisum sativum (garden pea)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999

C:Accession: T06477

R:Williamson, C.L.; Slocum, R.D.

Plant Physiol. 100, 1464-1470, 1992

A:Title: Molecular cloning and evidence for osmoregulation of the di-pyrroline-5-carboxylate reductase

A:Reference number: 215706

A:Accession: T06477

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-273 <M12>

A:Cross-references: EMBL:X62842; NID:g20850; PID:CAAA4646.1; PID:g20851

A:Experimental source: cv. Mando

C:Genetics:

A:Gene: PROC

C:Superfamily: pyrroline-5-carboxylate reductase

C:Keywords: oxidoreductase; proline biosynthesis

Query Match 30.6%; Score 479; DB 2; Length 273;

Best Local Similarity 40.1%; Pred. No. 9.3e-31;

Matches 108; Conservative 55; Mismatches 90; Indels 16; Gaps 5;

Db 2 SVGFTAGGLAVRF---TAAGILSAHKITA--SSEPMPLPVSA LRKMGVNLTSNKT 55

Db 12 TLGFTAGGMAESIAKAGARSGLVSSRIIVAHNSR---AAFEISGITVLSNDV 67

Db 56 VKHSDVFLAVPHIIPILDEIGADVQARHIVSCAAGVTISSEKKIMAFQAPKVR 115

Db 68 VRASNVVSVAPOLKVDVLTAKLPILTKDKLLVSAAGIKL---KDLQENAGHERFR 123

Db 116 CMTNPVVVQEGATVYATGTHALVDEGQLLEQLMSSVGFCTEVEDLIDAVTGLSGSPA 175

Db 124 VMPNPVPAAGVAAVSWSLCGAATEEDANISQLFGSIKIMKADKFFDALTGLSGSPA 183

Db 176 YAFMALD--ADGVKMGIPRRRLAIOGAQALLGAAKMLDSEQHPCOLKDNVCSPGAT 233

Db 184 YIYLALEALDGVAAAGLPRLDALSLASQTVGAASMATLSGRHGGQLKDDVTSFGTTI 243

Db 234 HALHFLSGGFRSLINAVEASCIRTEL 262

Db 244 AGVHELEKCGFRGTLTMNAVVAAKRSREL 272

RESULT 6

pyrroline-5-carboxylate reductase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000

C:Accession: T36286

R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1998

A:Reference number: 221603

A:Accession: T36286

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-284 <SEB>

A:Cross-references: EMBL:AL049819; PIDN:CA842663.1; GSPDB:GND0070; SCOEDB:SCET.04C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: PROC; SCOEDB:SCET.04C

C:Superfamily: pyrroline-5-carboxylate reductase

Query Match 27.0%; Score 423; DB 2; Length 284;

Best Local Similarity 40.1%; Pred. No. 2.8e-26;

Matches 108; Conservative 38; Mismatches 89; Indels 34; Gaps 5;

Db 4 GFITAGGLAVRF---TAAGILSAHKITA--SSEPMPLPVSA LRKMGVNLTR-----SNKET 55

Db 34 GMIGAG-----WAPADL-----VTARRRERADELAHRGVTPTVNAEA 72

Db 56 VKHSDVFLAVPHIIPILDEIGADVQARHIVSCAAGVTISSEKKIMAFQAPKVR 115

Db 73 AKADTLITLVKPDMDGLDELAPHPADRLVSGAAGVPTSFEEEL---APGTPVVR 129

Db 116 CMTNPVVVQEGATVYATGTHALVDEGQLLEQLMSSVGFCTEVEDLIDAVTGLSGSPA 175

Db 130 VMTNPVLDAMSVISAGTATTAHLTHEETGAVGKTRVPESSQODACTALSGSPA 189

Db 176 YAFMALD--ADGVKMGIPRRRLAIOGAQALLGAAKMLDSEQHPCOLKDNVCSPGAT 233

Db 190 YEFYLVAMTDAGILGLPRDKAHDLVQSAIGAAKMLRSGEHPVRLRENVTSFACCTI 249

Db 234 HALHFLSGGFRSLINAVEASCIRTEL 262

Db 250 NAIRELEHNGVRAALIALEAARDRSREL 278

RESULT 7

probable pyrroline-5-carboxylate reductase (EC 1.5.1.2) - Synecocystis sp. (strain P

C:Species: Synecocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S76767

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yae

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis

A:Reference number: S74322; MUID:97061201

A:Accession: S76767

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-267 <KAN>

A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BA18679.1; PID:g165

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Start codon: GTG

C:Superfamily: pyrroline-5-carboxylate reductase

C:Keywords: oxidoreductase; proline biosynthesis

Query Match 26.9%; Score 421; DB 2; Length 267;

Best Local Similarity 38.5%; Pred. No. 3.7e-26;

Matches 106; Conservative 47; Mismatches 96; Indels 26; Gaps 8;

Db 1 MSVGTAGGLAVRF---YRTAAGILSAHKITA--SSEPMPLPVSA LRKMGVNLTR-----SNKET 55

Db 3 IQLDITGGVAAEALRLAIEKRYAPEEITVGPB-----GAKRKYLOKTVQVRSPP 56

Db 52 NKEVVKHSDVFLAVPHIIPILDEIGADVQARHIVSCAAGVTISSEKKIMAFQAPKVR 111

Db 57 NOEAAVSEVLLAVKPDMDGLDELAPHPADRLVSGAAGVPTSFEEEL---APGTPVVR 129

Db 112 KVICMTNPVVVQEGATVYATGTHALVDEGQLLEQLMSSVGFCTEVEDLIDAVTGLSGSPA 175

Db 112 ATRAMPPTPATV--GAGMTAIAANKMVEPDQLAKAAITSAVGNVVEENMLDAVTGV 169

Db 170 SGSGPAAVAFMALD--ADGVKMGIPRRRLAIOGAQALLGAAKMLDSEQHPCOLKDNVCS 227

Db 170 SGSGPAAVVALMIEALADGVAAAGLPRLDALSLASQTVGAASMATLSGRHGGQLKDDVTSFGTTI 249

Db 228 PGCATTHALHFLSGGFRSLINAVEASCIRTEL 262

Db 230 PGCTTIAVAVLEKMGFRSAIIEAVRAAARSOEL 264

```

RESULT 8
H83533
pyrroline-5-carboxylate reductase [imported] - Escherichia coli (strain 0157:H7)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C:Accession: H83533
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Natter, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Natter, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
A:Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H83533
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-269 <STO>
A:Cross-references: GB:AE005174; NID:912513226; PIDN:AA654732.1; GSPDB:GN00145; UWGP:204
A:Experimental source: strain 0157:H7, substrain EDL933
C:Genetics:
A:Gene: proc
C:Superfamily: pyrroline-5-carboxylate reductase

Query Match 25.9%; Score 406.5; DB 2; Length 269;
Best Local Similarity 35.9%; Pred. No. 5.3e-25;
Matches 99; Conservative 51; Mismatches 107; Indels 19; Gaps 6;

OY 3 VGTGAGQLAFTTAAGILSAHKITASSPEMNLPT---VSALR-KMGVNLTRSKNEYKH 58
||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 IGFICGNGM-KAILGLIASGVLPQGIWYTPSPDKVAALHDFGINAESAQEVAGI 63

OY 59 SDVFLAVKPHIIFILDEIGADVQARHIVSCAGVTSSEKKLMFQAPKVIKMT 118
||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 ADITFAVKPGIMIVLEISLTKNDLVYSIAAGVTLIDOLAR---ALGHRKRIIRAMP 120

OY 119 NTPVVOEGATVYATGTALVEDGQLLEQLMSSVGCTEVEDLIDAVTGLSGSPAYAF 178
||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 NTPALVNGMTSVTPNALVTPEDTADVLNIFRCFGEAEVIAEPMIHVYGVSGSPAYAF 180

OY 179 MALD--ADGVKMGILPRRLAIOIGAALLGAAMLLDSQHPCQKDNVCSPGGATIHAL 236
||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 MFEEMAADAIVLGMPRAQAKFPAQAQVWGSAMVLETEGHPALKDWCSPGGTTEAV 240

OY 237 HFLESGFRSLINAVEASCIPTRELQSMADQEKIS 272
||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 RVLEKGFRAVITEAM-----TKMEKSEKLS 267

RESULT 9
RDECC
pyrroline-5-carboxylate reductase (EC 1.5.1.2) - Escherichia coli
C:Species: Escherichia coli
C:Date: 13-Jun-1993 #sequence_revision 13-Jun-1993 #text_change 11-Jun-1999
C:Accession: A00385; B64767
C:Deutch, A.H.; Smith, C.J.; Rushlow, K.E.; Kretscher, P.J.
Nucleic Acids Res. 10, 7701-7714, 1982
A:Title: Escherichia coli deltal(1)-pyrroline-5-carboxylate reductase: gene sequence, ptc
A:Reference number: A00385; MUID:8315986
A:Accession: A00385
A:Molecule type: DNA
A:Residues: 1-269 <DEU>
A:Cross-references: GB:J01665; NID:9147358; PIDN:AA86433.1; PID:9147359
A>Note: parts of this sequence, including the amino and carboxyl ends of the mature prot
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:9742617
A:Accession: B64767
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-269 <BLAT>
A:Cross-references: GB:AE000145; GB:U00096; NID:91786580; PIDN:AA073489.1; PID:91786585;
A:Experimental source: strain K-12, substrain MG1655

```

```

C:Genetics:
A:Gene: proc
A:Map position: 9 min
C:Function:
A:Description: catalyzes reduction of pyrroline-5-carboxylate to proline
A:Pathway: proline biosynthesis
A>Note: third enzyme in the proline biosynthetic pathway
C:Superfamily: pyrroline-5-carboxylate reductase
C:Keywords: oxidoreductase; proline biosynthesis

Query Match 25.7%; Score 402.5; DB 1; Length 269;
Best Local Similarity 35.5%; Pred. No. 1.1e-24;
Matches 98; Conservative 52; Mismatches 107; Indels 19; Gaps 6;

OY 3 VGTGAGQLAFTTAAGILSAHKITASSPEMNLPT---VSALR-KMGVNLTRSKNEYKH 58
||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 IGFICGNGM-KAILGLIASGVLPQGIWYTPSPDKVAALHDFGINAESAQEVAGI 63

OY 59 SDVFLAVKPHIIFILDEIGADVQARHIVSCAGVTSSEKKLMFQAPKVIKMT 118
||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 ADITFAVKPGIMIVLEISLTKNDLVYSIAAGVTLIDOLAR---ALGHRKRIIRAMP 120

OY 119 NTPVVOEGATVYATGTALVEDGQLLEQLMSSVGCTEVEDLIDAVTGLSGSPAYAF 178
||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 NTPALVNGMTSVTPNALVTPEDTADVLNIFRCFGEAEVIAEPMIHVYGVSGSPAYAF 180

OY 179 MALD--ADGVKMGILPRRLAIOIGAALLGAAMLLDSQHPCQKDNVCSPGGATIHAL 236
||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 MFEEMAADAIVLGMPRAQAKFPAQAQVWGSAMVLETEGHPALKDWCSPGGTTEAV 240

OY 237 HFLESGFRSLINAVEASCIPTRELQSMADQEKIS 272
||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 RVLEKGFRAVITEAM-----TKMEKSEKLS 267

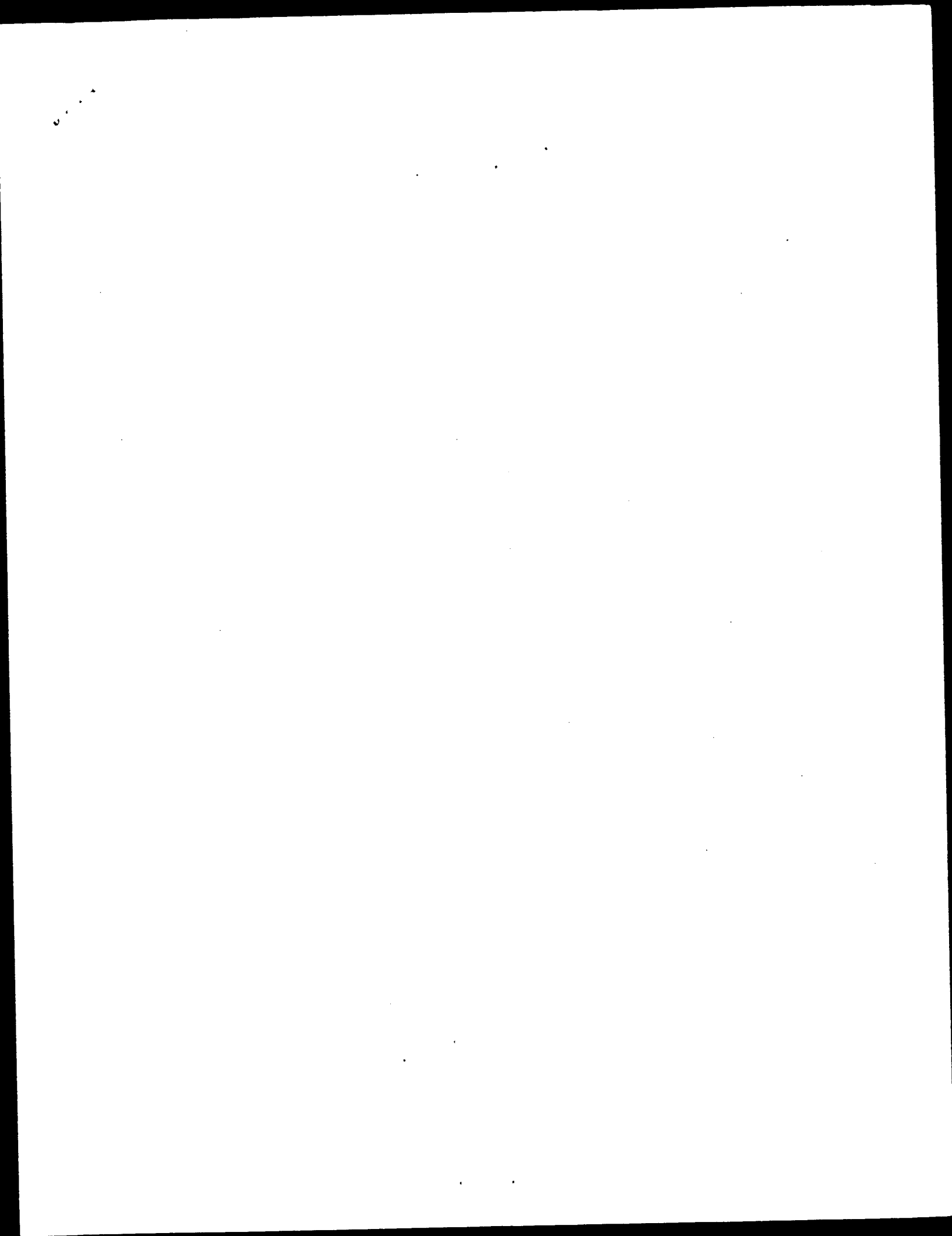
RESULT 10
JQ0418
pyrroline-5-carboxylate reductase (EC 1.5.1.2) precursor - Pseudomonas aeruginosa (st
C:Species: Pseudomonas aeruginosa
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 31-Dec-2000
C:Accession: JQ0418; H83597
R:Savioz, A.; Jeenes, D.J.; Kocher, H.P.; Haas, D.
Gene 86, 107-111, 1990
A:Title: Comparison of proc and other housekeeping genes of Pseudomonas aeruginosa w
A:Reference number: JQ0418; MUID:90185238
A:Accession: JQ0418
A:Molecule type: DNA
A:Residues: 1-273 <SAV>
A:Cross-references: GB:M33557; NID:9151509; PIDN:AAA25975.1; PID:9151510
A:Experimental source: strain PA01
A>Note: part of this sequence, including the amino end of the mature protein, was con
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; L
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: H83597
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-273 <STO>
A:Cross-references: GB:AE004476; GB:AE004091; NID:99946234; PIDN:AA030782.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: proc; PA0393
C:Superfamily: pyrroline-5-carboxylate reductase
C:Keywords: oxidoreductase; proline biosynthesis
F:2-273/Product: pyrroline 5-carboxylate reductase #status experimental <MAT>

Query Match 24.1%; Score 377; DB 2; Length 273;
Best Local Similarity 36.6%; Pred. No. 1.2e-22;

```

Db	19	VAFIAGSMA-EGMISGIVRANKIPONICVTNRSNTERLELELOGIGKALPNOLCIE	77
QY	58	HSDFVLAVKPHIIPFLIDEIGADVOARHIVYSCAAGVTISSEYKKLMAFOPAPVIRCM	11
Db	78	DMVDVILAMKPKAENLNLSTLSKRSRQIPHOLISVLAIGTTSFLEOSTLNBP---	13
QY	118	TNTPVVVOEGATVYVAGTHALVEDCQLEQLMSVSGCTEVEBDLIDAVTGLSGSGPAPA	177
Db	135	PMTSMIGASATVIALGKYSEDLKKLAEALIGCGGEVYTIQENQMDIFTTGIAGSGPAYF	194
QY	178	FMAUD--ADGGVMMKJPRRLAIOLOAQAALLLSAANKMLDSQHCQKDNVCSGGCATTHA	235
Db	195	YLMDEFLEKTEGEELGDLKOLSHSIGAQTLLGAARKMLMETGEHPEILRDNITSPNGTTAAG	254
QY	236	LHFL-ESGGRSLI	248
Db	255	LQALKKSGEKRL	268
RESULT	12		
	572897	pyrroline-5-carboxylate reductase (EC 1.5.1.2) - Mycobacterium leprae	
	N:Alternate names:	B2168_C2.211 protein	
	C:Species:	Mycobacterium leprae	
	C:Date:	19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2000	
	C:Accession:	S72897	
	R:Smith, D. R.;	Robison, K.	
	A:Description:	the EMBL Data Library, November 1993	
	A:Reference number:	S72866	
	A:Accession:	S72897	
	A:Status:	preliminary	
	A:Molecule type:	DNA	
	A:Residues:	1-294 <SMI>	
	A:Cross-references:	EMBL:U00018; NID:g467037; PIDN:AAAI7253.1; PID:g467049	
	C:Genetics:		
	A:Gene:	proc	
	C:Superfamily:	pyrroline-5-carboxylate reductase	
	C:Keywords:	oxidoreductase	
Query Match		20.88; Score 326.5; DB 2; Length 294;	
Best Local Similarity		36.28; Pred. No. 1.4e-18;	
Matches	88; Conservative	45; Mismatches 77; Indels 33; Gaps	
QY	48	LTRSNKETVAKHSDVFLAVKPHIIPFLIDEI-----GADVOARHIVYSCAAGVTISSEK	102
Db	54	LTVSYVDVAVENMFEVYVAVKTTDVESWGDLYOAAVAVNADSAEQLVTVVAGVITYLTS	113
QY	103	KLMAFOPAPVIRCMTNTPVVVOEGATVYVAGTHALVEDCQLEQLMSVSGCTEVE	159
Db	114	KLPAGTP---VVRAMPNAALVGVAGTVYLAKEGFV---TGOQFEVLYAMEDVAGGVLYTP	167
QY	160	EDLIDAVYGLSGSGPAPAFMALDA--DGGYKMGJPRRLAIOLOAQAALLGAANKML--DS	214
Db	168	ESQMDAVYAVSGSPAYFFLVELVADAGVAVGLTRQVATLTAQTGMSAAMLLRMDQ	227
QY	215	EQH-----PCOLKDNVCSGGCATTHALHFLFESGGRSLIINAVASCTIRP	260
Db	228	DRHSAEVAPLCAQVDVPAQAQRLRTTISPGTGTAAALRELEGRGLRMVYDAVAQAARKIRSE	287
QY	261	EQI	263
Db	288	QLR	290

RESULT 13
 PYRROLINE-5-CARBOXYLATE REDUCTASE homolog yqj0 - *Bacillus subtilis*
 N:contains: oxidoreductase (EC 1.-.-.-)
 C:Species: *Bacillus subtilis*
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C:accession: G69964



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:53:55 ; Search time 24.93 seconds

(without alignments)
1842.339 Million cell updates/sec

Title: US-09-912-717-1

Sequence: 1 MSVGFAGAGLAFRTAAGT.....TPSSPKLLTRSLALGKKMD 314

Scoring table: BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_17:*
2: SP_archaea:*
3: SP_bacteria:*
4: SP_fungi:*
5: SP_human:*
6: SP_invertebrate:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1231.5	78.6	319	4	O9Y5J4 homo sapien
2	898	57.3	272	4	O9H8Q4
3	578	36.9	279	5	O21544
4	537.5	34.3	274	4	O9H896
5	522.5	33.3	274	11	O9DCC4
6	503	32.1	273	5	O9VEJ3
7	500.5	31.9	274	10	O9AVM3
8	481	30.7	280	5	O9V3F8
9	469	29.9	280	5	O96643
10	456	29.1	270	1	O9H899
11	448	28.6	266	2	O87725
12	443	27.0	284	2	O9X8G1
13	399	25.5	272	2	O9K308
14	380	24.3	248	11	O9DOX2
15	376	24.0	254	5	O9NGS4
16	344	22.0	279	1	O9H8T6
17	325.5	20.8	273	1	O9H8T6
18	323.5	20.6	279	2	O9KCR6
19	322	20.5	209	5	O27722

20	319	20.4	255	2	O9W247	O9W247 thermotoga
21	301	19.2	290	2	O9E283	O9E283 bradyrhizob
22	297.5	19.0	305	2	O9PA08	O9PA08 xyloella fas
23	295.5	18.9	262	2	O9CEP7	O9CEP7 lactococcus
24	293	18.7	266	1	O9YFV4	O9YFV4 aeropyrum p
25	292.5	18.7	282	3	O9P777	O9P777 schizosach
26	289	18.5	275	2	O9CPE8	O9CPE8 pasteurella
27	289	18.4	264	2	O9RUV7	O9RUV7 delnoccocus
28	263.5	16.8	271	2	O9FDK7	O9FDK7 zymomonas m
29	258.5	16.5	260	2	O9AAU9	O9AAU9 caulobacter
30	258.5	16.5	272	2	O9KUD5	O9KUD5 vibrio chol
31	258	16.5	256	2	O9AIS9	O9AIS9 streptococc
32	252.5	16.1	271	2	O9RTZ0	O9RTZ0 staphylococ
33	229	14.6	267	2	O9KEG5	O9KEG5 bacillus ha
34	212.5	13.6	263	2	O9K1N1	O9K1N1 neisseria m
35	212.5	13.6	263	2	O9JWM2	O9JWM2 campylobact
36	192.5	12.3	243	2	O9PNL8	O9PNL8 campylobact
37	179	11.4	275	2	O9KDB6	O9KDB6 bacillus ha
38	151.5	9.7	117	10	O41907	O41907 arabidopsis
39	151	9.6	99	2	O07508	O07508 bacillus su
40	148.5	9.5	253	1	O26983	O26983 methanobact
41	114.5	7.3	264	1	O27779	O27779 methanobact
42	106.5	6.8	250	4	O95435	O95435 homo sapien
43	106.5	6.8	1043	2	O910V6	O910V6 pseudomonas
44	101.5	6.5	350	13	O98S39	O98S39 salmo salar
45	101	6.4	8817	2	O53840	O53840 polyangium

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	319 AA.
O9Y5J4				
ID O9Y5J4				
AC O9Y5J4				
DT 01-NOV-1999 (TREMURel. 12, Created)				
DT 01-NOV-1999 (TREMURel. 12, Last sequence update)				
DT 01-JUN-2001 (TREMURel. 17, Last annotation update)				
DE PYROLINE 5-CARBOXYLATE REDUCTASE ISOFORM.				
OS P5CR2.				
GN Homo sapiens (human).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX NCBI_TaxID:9606;				
RN [1]				
RP SEQUENCE FROM N.A.				
RA Kok L.D.S., Au T.C.C., Fung K.P., Tsui S.K.W., Lee C.Y., Waye M.M.Y.;				
RT "Molecular characterization of a novel human pyroline-5-carboxylate				
RT reductase isolated in liver."				
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.				
DR EMBL: AF151351; AAD34611.1; -				
DR InterPro: IPR000304; P5CR.				
DR Pfam: PF01089; P5CR; 1.				
DR PROSITE: PS00521; P5CR; 1.				
SQ SEQUENCE 319 AA; 33726 MW; 01D497002A598838 CRC64;				

Query Match 78.6%; Score 1231.5; DB 4; Length 319;

Best Local Similarity 81.7%; Pred. No. 3.8e-90; Mismatches 41; Indels 11; Gaps 4;

OY	1	MSVGFAGAGLAFRTAAGTSAHKT-IASSPEMLPTVSA LRKGVLTGRSNEFKHS 59	
DB	1	MSVGFAGAGLAFRTAAGTSAHKT-IASSPEMLPTVSA LRKGVLTGRSNEFKHS 59	
OY	60	D-----VLELAVKPHIIPILDEIGADVOARHYVSCAGVTISSVEKKLMAOPARKVI 114	
DB	58	DEDAQRPRPVSCEAHIIIPILVEIGADVOARHYVSCAGVTISSVEKKLMAOPARKVI 117	
OY	115	RQMTNPVYVVGATVATGTHALVEDGQLLEQLMSSVGFCTEVEDLITAVYGLSGSP 174	
DB	118	RQMTNPVYVVGATVATGTHALVEDGQLLEQLMSSVGFCTEVEDLITAVYGLSGSKP 177	

```

OY 175 AYAFMAMD--ADGVKMGDLPRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 232
DB 178 AYAFMAMDALADGVKMGDLPRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 237
OY 233 IHALHFLSEGSFRLINAVASCIPTRELQSMADQEKISPAALKTLLDRVYLESPYVS 292
DB 238 IHALHFLSEGSFRLINAVASCIPTRELQSMADQEKISPAALKTLLDRVYLESPYVS 297
OY 293 TLTPSSPGKLLTRSLALGKKD 314
DB 298 TLTPSSPGKLLTRSLALGKKD 319

RESULT 2
OY 09HB04 PRELIMINARY; PRT: 272 AA.
AC 09HB04;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHETICAL 28.5 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gu J.R., Man D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Gu W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel Human cDNA clones with function of inhibiting cancer cell
RT growth.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF218000; AAG17242.1; -.
DR InterPro: IPR000304; P5CR.
DR Pfam: PF01089; P5CR; 1.
KW Hypothetical protein.
SQ SEQUENCE 272 AA; 28471 MW; 0BC71EA751F78F4 CRC64;

Query Match 57.3%; Score 898; DB 4; Length 272;
Best Local Similarity 83.7%; Pred. No. 1e-63;
Matches 180; Conservative 19; Mismatches 10; Indels 6; Gaps 2;

OY 1 MSVGFAGAGOLAYR----FTAGILSAHKTIASSPEMNLPTYSALPKMGVNLTRSKETV 56
DB 1 MSVGFAGAGOLAYR----FTAGILSAHKTIASSPEMNLPTYSALPKMGVNLTRSKETV 60
OY 57 KHSVDFLAVKPHIIPFLDEIGADVOARHIVVSCAAGVITSSVEKKLMAFOPAPKVIK 116
DB 61 QHSDVFLAVKPHIIPFLDEIGADIDRHIVVSCAAGVITSSVEKKLSAFRAPRVIRC 120
OY 117 MNTNPVVVQEGATVATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGSPAY 176
DB 121 MNTNPVVVQEGATVATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGSPAY 180
OY 177 AFMAMD--ADGVKMGDLPRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 209
DB 181 AFMAMD--ADGVKMGDLPRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 215

RESULT 3
OY 021544 PRELIMINARY; PRT: 279 AA.
AC 021544;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE M153x1 PROTEIN.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.

```

```

OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews P.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kistner J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smailson N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Spratt J., Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL: Z67995; CAB91943.1; -.
DR InterPro: IPR000304; P5CR.
DR Pfam: PF01089; P5CR; 1.
SQ SEQUENCE 279 AA; 29165 MW; C68CAD4BC4EB20E6 CRC64;

Query Match 36.9%; Score 578; DB 5; Length 279;
Best Local Similarity 44.6%; Pred. No. 3.2e-38;
Matches 127; Conservative 55; Mismatches 89; Indels 14; Gaps 4;

```

```

OY 1 MSVGFAGAGOLAYR----AGILSAHKTIASSPEMNLPTYSALPKMGVNLTRSKETV 56
DB 1 MKRIGFAGAGOLAYR----AGILSAHKTIASSPEMNLPTYSALPKMGVNLTRSKETV 60
OY 57 KHSVDFLAVKPHIIPFLDEIGADVOARHIVVSCAAGVITSSVEKKLMAFOPAPKVIK 116
DB 61 QHSDVFLAVKPHIIPFLDEIGADIDRHIVVSCAAGVITSSVEKKLSAFRAPRVIRC 117
OY 117 MNTNPVVVQEGATVATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGSPAY 176
DB 118 MNTNPVVVQEGATVATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSSGSPAY 177
OY 177 AFMAMD--ADGVKMGDLPRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 234
DB 178 MFAVIEGLADGVKMGDLPRRLAIOGAQALLGAAMKMLDSEPHCOLKDNVCSGSGAT 237
OY 235 ALHFLSEGSFRLINAVASCIPTRELQSMADQEKISPAALKT 279
DB 238 GMHFLSEGSFRLINAVASCIPTRELQSMADQEKISPAALKT 277

RESULT 4
OY 09H896 PRELIMINARY; PRT: 274 AA.
AC 09H896;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CDNA FLJ13852 FTS, CLONE THYR0100934, WEAKLY SIMILAR TO PYRROLINE-5-
DE CARBOXYLATE REDUCTASE (EC 1.5.1.2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=THYROID GLAND;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nadekura T., Ishii S., Kawai Y., Saito Y., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y., Oshima A.;

```

"MEDD human CDNA sequencing project."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK023914; BAB14721.1; -
 DR InterPro: IPR000304; P5CR.
 DR Pfam: PF01089; P5CR.1.
 DR PROSITE: PS00521; P5CR.1.
 SQ SEQUENCE 274 AA; 28663 MW; 846FDEC603F3B548 CRC64;

Query Match 34.3%; Score 537.5; DB 4; Length 274;
 Best Local Similarity 45.3%; Pred. No. 5; Le-35;
 Matches 121; Conservative 45; Mismatches 88; Indels 13; Gaps 5;

QY 3 VGFAGAGCQ---LAVFTAGILSAHKIIASSP-EMNLPTVSAIRKMGVLTIRSNKTVK 57
 Db 11 VGFVAGAGMAGAIAGGLIRACKVKAQHILASAPIDRLNCHQAL---GCRTHSNQEVLD 67
 QY 58 HSDVFLAVKPHIIPFLIDIGADVQARHIVSCAGVTTSSVEKKIMAFQAPKVIKCM 117
 Db 68 SCLVIFATKPHVLPVLAELVAPVTEHILVSAVAGSLSTLELL---PPNTRVLRLV 124
 QY 118 TNPVVOEGATVATGTHALVEDGQLLEDMSSVGFCTVEEDLIDAVTGLSGSPAY- 176
 Db 125 PNLPCVVOEGAVIMARGNHVSSEFKLQHLLEACGRPEEVPAYVVDIHTGLSGSVAFV 184
 QY 177 AFMALDADGCVKMGKLPRLAIOGAQALLGAARMLDSEQHPCQLKDNVCSFGATTHA 235
 Db 185 CAFSALAEAGAVKMGKMSLAHRLAQTLLCTAKMLHEGHPAQDLSDVCTPGCTTIYG 244
 QY 236 LHFLESGFRSLINAVEASCITREL 262
 Db 245 LHALEGGGLRAATWSAVEAATCRAKEL 271

RESULT 5
 ID 09DCC4 PRELIMINARY; PRT; 274 AA.
 AC 09DCC4;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE 1110058B13RIK PROTEIN.
 GN 1110058B13RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner I., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Botfield D., Boujuna N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustinglich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Saeki H., Sato K., Schoenbach C., Seye T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontsuki S.,
 Hayashizaki Y.;
 RA *Functional annotation of a full-length mouse CDNA collection.*;
 RT Nature 409:685-690(2001).
 RL EMBL: AK002912; BAB2451.1; -
 DR MGI:1913444; 1110058B13RIK.

DR InterPro: IPR000304; P5CR.
 DR Pfam: PF01089; P5CR.1.
 SQ SEQUENCE 274 AA; 28694 MW; FE3892C01C6068A5 CRC64;

Query Match 33.3%; Score 522.5; DB 11; Length 274;
 Best Local Similarity 42.9%; Pred. No. 6; Le-34;
 Matches 114; Conservative 47; Mismatches 94; Indels 11; Gaps 4;

QY 3 VGFAGAGCQ---LAVFTAGILSAHKIIASSP-EMNLPTVSAIRKMGVLTIRSNKTVK 58
 Db 11 VGFVAGAGMAGAIAGGLIRACKVKAQHILASAPIDRLNCHQAL---GCRTHSNQEVLD 68
 QY 59 SDVFLAVKPHIIPFLIDIGADVQARHIVSCAGVTTSSVEKKIMAFQAPKVIKCM 118
 Db 69 SCLVIFATKPHVLPVLAELVAPVTEHILVSAVAGSLSTLELL---PPNTRVLRLV 125
 QY 119 NTPVVOEGATVATGTHALVEDGQLLEDMSSVGFCTVEEDLIDAVTGLSGSPAY- 176
 Db 126 PNLPCVVOEGAVIMARGNHVSSEFKLQHLLEACGRPEEVPAYVVDIHTGLSGSVAFV 185
 QY 177 AFMALDADGCVKMGKLPRLAIOGAQALLGAARMLDSEQHPCQLKDNVCSFGATTHA 236
 Db 186 TFSALAEAGAVKMGKMSLAHRLAQTLLCTAKMLHEGHPAQDLSDVCTPGCTTIYG 245
 QY 237 LHFLESGFRSLINAVEASCITREL 262
 Db 246 LHALEGGGLRAATWSAVEAATCRAKEL 271

RESULT 6
 ID 09VEJ3 PRELIMINARY; PRT; 273 AA.
 AC 09VEJ3;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE CG5840 PROTEIN.
 GN CG5840.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyroidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Byrne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins J., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner J., Henderson S.N.,
 Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Anril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
 Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Borkman M.R., Bouck J., Brokstein P., Brotler P.,
 Burris K.C., Busam S.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Paalos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
 Jallali M., Kalush F., Kapten G.H., Ke Z., Kenison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Matvei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

Query Match	31.98;	Score 500.5;	DB 10;	Length 274;
Best Local Similarity	41.58;	Pred. No. 4.5e-32;		
Matches 112;	Conservative 51;	Mismatches 90;	Indels 17;	Gaps 5

RA Spitzkas R., Operating Officer, Department of Engineering, Sun D.
RA Spitzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Massaman D.A., Weinsbach G.M., Weissenbach J.,
RA Williams S.M., Woodgate T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN 121

RP SEQUENCE FROM N.A.

RC STRAIN-CS:

RA Miesner S.R., Walker V.K.;

RT "High density of unrelated genes showing overlapping and intraltronic
 RT transcribed units in *Drosophila*.";
 RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.

DR EMBL: A6003724; AAF55626.1; -
 DR EMBL: AF170829; AAD49740.1; -
 DR FlyBase: FBgn0015781; PSCR.

DR InterPro: IPR000304; PSCR.

DR Pfam: PF01089; PSCR; 1.
 DR PROSITE: PS00521; PSCR; 1.
 DR Oxidoreductase.

SO SEQUENCE 280 AA; 29621 MW; D3D4PD9611574684 CMC64;

Query Match

Best Local Similarity 30.7%; Score 481; DB 5; Length 280;
 Matches 108; Conservative 51; Mismatches 95; Indels 22; Gaps 6;

QY 3 VGFIGAGQLAYR---FTAAGILSAHKIIASSPEKMLPTVSALKMGVNLTRSKNETVKH 58
 DB 8 IGFIGGNMVAIGSLVRCGIVKASQVQVSP--HIEMLQWRDLCAVTCDDNCVLEH 65
 QY 59 SDVFLAVKPHIIPILDEIGADVQARHI-----VSCAGAVTISVEKKLMFOP 109
 DB 66 SDIVFLCAKPH---MLTPCAQLKYYKHVPASAKDASKLVSLVLAQTSLETFLEAFSPFGS 121
 QY 110 AP-KVIRCMNTPPVVOEGATVATGTHALVEDGQLLEQLMSSVGFCTVEEDLIDAVTG 168
 DB 122 SELKVIKRTMPTNSMVGECCTVYTGNAVSHHDEKTHLMLGALQVPESMIDAVTG 181
 QY 169 LSGSPAYAFMALD--ADGVKMGLPRLAIQLGAQLLGAARKMLDSEQHPQLKDNVC 226
 DB 182 VAGCGFAFYTTIEALADGVKQGVPRQMALQFAAQTLLGAATVLLTGKHPAVLREVC 241
 QY 227 SPGATIHALHFLSEGGFRSLINAVEASCIRREL 262
 DB 242 SPGATIVGVHELEKGNLRSTLINAVEKSSORSABL 277

RESULT 9

ID 096643

AC 096643; PRELIMINARY; PRT; 280 AA.

DT 01-MAY-1999 (Tremblrel. 10, Created)

DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)

DE PYRROLINE-5-CARBOXYLATE REDUCTASE.

OS *Drosophila melanogaster* (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Ephydroidea; Diptera; Brachycera; Muscomorpha;

NCBI_Taxid=7227;

RN 111

RP SEQUENCE FROM N.A.

RC STRAIN-CS:

RA Miesner S.R., Walker V.K.;

RT "Cloning and sequence analysis of the gene encoding pyrroline 5-
 RT carboxylate reductase from *Drosophila melanogaster*.";
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF098020; AAC70780.1; -
 DR FlyBase: FBgn0015781; PSCR.
 DR InterPro: IPR000304; PSCR.
 DR Pfam: PF01089; PSCR; 1.
 DR PROSITE: PS00521; PSCR; 1.
 SO SEQUENCE 280 AA; 29688 MW; 2A7416D39D2A2369 CRC64;

Query Match 29.9%; Score 469; DB 5; Length 280;
 Best Local Similarity 38.4%; Pred. No. 1.5e-29;
 Matches 106; Conservative 51; Mismatches 97; Indels 22; Gaps 6;

QY 3 VGFIGAGQLAYR---FTAAGILSAHKIIASSPEKMLPTVSALKMGVNLTRSKNETVKH 58
 DB 8 IGFIGGNMVAIGSLVRCGIVKASQVQVSP--HIEMLQWRDLCAVTCDDNCVLEH 65
 QY 59 SDVFLAVKPHIIPILDEIGADVQARHI-----VSCAGAVTISVEKKLMFOP 109
 DB 66 SDIVFLCAKPH---MLTPCAQLKYYKHVPASAKDASKLVSLVLAQTSLETFLEAFSPFGS 121
 QY 110 AP-KVIRCMNTPPVVOEGATVATGTHALVEDGQLLEQLMSSVGFCTVEEDLIDAVTG 168
 DB 122 SELKVIKRTMPTNSMVGECCTVYTGNAVSHHDEKTHLMLGALQVPESMIDAVTG 181
 QY 169 LSGSPAYAFMALD--ADGVKMGLPRLAIQLGAQLLGAARKMLDSEQHPQLKDNVC 226
 DB 182 VAGCGFAFYTTIEALADGVKQGVPRQMALQFAAQTLLGAATVLLTGKHPAVLREVC 241
 QY 227 SPGATIHALHFLSEGGFRSLINAVEASCIRREL 262
 DB 242 SPGATIVGVHELEKGNLRSTLINAVEKSSORSABL 277

RESULT 10

ID 09HH99

AC 09HH99; PRELIMINARY; PRT; 270 AA.

DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)

OS *Methanosarcina acetivorans*.

OC Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;

NCBI_Taxid=2214;

RN 111

RP SEQUENCE FROM N.A.

RC STRAIN-CS:

RA Zhang J.K., White A.K., Kuetner H.C., Boccazzi P., Metcalf W.W.;

RT "Molecular and genetic analysis of proline biosynthesis in the
 RT methanogenic archaeon *Methanosarcina acetivorans* C2A.";
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF305580; AAG22033.1; -
 DR InterPro: IPR000304; PSCR.
 DR Pfam: PF01089; PSCR; 1.
 DR PROSITE: PS00521; PSCR; 1.
 SO SEQUENCE 270 AA; 27943 MW; 50EC656AFC10B1CF CMC64;

Query Match 29.1%; Score 456; DB 1; Length 270;
 Best Local Similarity 39.7%; Pred. No. 1.5e-28;
 Matches 106; Conservative 52; Mismatches 97; Indels 12; Gaps 5;

QY 3 VGFIGAGQLAYR---FTAAGILSAHKIIASSPEKMLPTVSALKMGVNLTRSKNETVKH 57
 DB 6 IGFIGAGKMSALMOGTIKAGIVTPENIGAS--DYEPFLADLQKIGIRSTDNVAIVR 63
 QY 58 HSDVFLAVKPHIIPILDEIGADVQARHIIVSCAGAVTISVEKKLMFOPAPKVRICM 117
 DB 64 ESDILFLAVKPHIIPILDEIGADVQARHIIVSCAGAVTISVEKKLMFOPAPKVRICM 120
 QY 118 TMTPPVVOEGATVATGTHALVEDGQLLEQLMSSVGFCTVEEDLIDAVTGLSGSPAYA 177
 DB 121 PNIATVSEPAASGIAIPGKNAITPBDKALAEIFSAGVAVQVPESLMDAVTGLSGSPAFI 180
 QY 178 FMALD--ADGVKMGLPRLAIQLGAQLLGAARKMLDSEQHPQLKDNVCSPGATITIA 235
 DB 181 FVITAMADGAVLBSGMRKSLALTGAOTVLGAARKMALETGMHNGELKMDVTSAGTTIGG 240

QY 236 LHFLESGFRSLINAVEASCIRREL 262
 DB 241 IHSLEAGIRAAFMNAVIRASERSKEL 267

RESULT 11
 087725 PRELIMINARY; PRT; 266 AA.

AC 087725;
 DT 01-NOV-1998 (TREMUREL. 08, Created)
 DT 01-NOV-1998 (TREMUREL. 08, Last sequence update)
 DT 01-JUN-2001 (TREMUREL. 17, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2).
 GN PROC.
 OS Clostridium sticklandii.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OC NCBI_TaxID=1511;
 RN [1]
 RC STRAIN-HF, DSM 517;
 RA Kenkies J., Ziehm R., Fritzsche K., Andreesen J.R., Pich A.;
 RT "Proline biosynthesis in Clostridium sticklandii: Purification of
 RT pyrroline-5-carboxylate reductase and characterization of the encoding
 RT gene proc.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ010739; CA09332.1; .
 DR InterPro; IPR000304; PSCR.
 DR Pfam; PF01089; PSCR; 1.
 DR PROSITE; PS00521; PSCR; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 266 AA; 28902 MW; 48B5AD46A5A43B65 CRC64;

Query Match 28.6%; Score 448; DB 2; Length 266;
 Best Local Similarity 38.5%; Pred. No. 6.5e-28;
 Matches 104; Conservative 51; Mismatches 97; Indels 18; Gaps 5;

QY 3 VGFIGAGOLAYRPTAA-----GILSAHKIIASPEMNLPTVSALRKM-----GVNLTRSKKE 54
 DB 5 IGTIGCGNMAQAMISALVSKLIESNQIIVSNRSKNI-----LEKMNKEYGITTIANNIE 59

QY 55 TVKHSVLFYAVPHIIFLIDETIGADVOARHIVVSCAGVYTSSEKKIMAFOPAPYIR 114
 DB 60 VAEKCDIVFLAVKONYEMVWKIKDSYTKDKLEVSIAQKTMETLEHLAGY---AKIL 116

QY 115 RCMNTPVVOEGATVYATGTHALVEDGOLLLEOLMSSVGFCTEVEDLIDAVTGLSGGP 174
 DB 117 RTMNTPTSMVSEGMTAICPNSNIASEBELVLTIESFGAVETIEEKLFDVAVAVSGSSP 176

QY 175 AYAFMALD--ADGVKMGDLPRRLAIOGAQALLGAAKMLDSBOHCQKLDKDVCSFGGAT 232
 DB 177 AYVFMFIEAMADAIVIOGMRAQYRPAQAVVGSAAKMLVSKSHPEGLKDVCSFGGTT 236

QY 233 IHALHFLSESGFRSLINAVEASCIRREL 262
 DB 237 IEAVALVEERGMRSVAIVLEAMKCKQKSKEM 266

RESULT 12
 09X8G1 PRELIMINARY; PRT; 284 AA.

AC 09X8G1;
 DT 01-NOV-1999 (TREMUREL. 12, Created)
 DT 01-NOV-1999 (TREMUREL. 12, Last sequence update)
 DT 01-JUN-2001 (TREMUREL. 17, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE.
 GN SCE7.04C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=1902;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Seeger K.J., Harris D.;
 RT "A set of ordered cosmid and a detailed genetic and physical map for
 RT the 8 Mb streptomycetes coelicolor A3(2) chromosome.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb streptomycetes coelicolor A3(2) chromosome.";
 RL M.O. Microbiol. 21:77-96(1996).
 DR EMBL; AL049819; CA842663.1; .
 DR InterPro; IPR000304; PSCR.
 DR Pfam; PF01089; PSCR; 1.
 DR PROSITE; PS00521; PSCR; 1.
 SQ SEQUENCE 284 AA; 29806 MW; 291D9899EB4AB115 CRC64;

Query Match 27.0%; Score 423; DB 2; Length 284;
 Best Local Similarity 40.1%; Pred. No. 7e-26;
 Matches 108; Conservative 38; Mismatches 89; Indels 34; Gaps 5;

QY 4 GFIGAGOLAYRPTAAGILSAHKIIASPEMNLPTVSALRKMGLVLT-----SKKET 55
 DB 34 CMIGAG-----WAPADLT-----VTARRRERAEDELARRHGVPTVTAEA 72

QY 56 VKHSVLFYAVPHIIFLIDETIGADVOARHIVVSCAGVYTSSEKKIMAFOPAPYIR 115
 DB 73 AKAADPLITVPRQDMGTLDLAVHPADRLVIGAGVYTSFEERL---AGTPEYVR 129

QY 116 CMNTPVVOEGATVYATGTHALVEDGOLLLEOLMSSVGFCTEVEDLIDAVTGLSGGPA 175
 DB 130 VMNTPTALVDENASVTSAGTHATAHLTETETIFGAVKTLRVPSQDCTALSGSGPA 189

QY 176 YAFMALDA--DGVKMGDLPRRLAIOGAQALLGAAKMLDSBOHCQKLDKDVCSFGGAT 233
 DB 190 YFFYLVEAMTDAGILGLPRDKAHDLIVOSAIKAAMLRDSEHVKLRBNTVSPAGTTI 249

QY 234 HALHFLSESGFRSLINAVEASCIRREL 262
 DB 250 NATRELENHGVRAALIAALEARDRSREL 278

RESULT 13
 09K308 PRELIMINARY; PRT; 272 AA.

AC 09K308;
 DT 01-OCT-2000 (TREMUREL. 15, Created)
 DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)
 DT 01-JUN-2001 (TREMUREL. 17, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2).
 GN PROC.
 OS Bacillus cereus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OC NCBI_TaxID=1596;
 RN [1]
 RC STRAIN-ATCC 14579;
 RA Okstad O.A., Agasse H., Iereclus D., Kolsto A.B.;
 RT "20 kb DNA sequence surrounding the hblA hemolytic enterotoxin locus
 RT from Bacillus cereus ATCC 14579.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

Thu Jan 31 15:07:59 2002

us-09-912-717-1.rpt

Page 8

Db 110 EGYTSTISGNSAVT PDDDEAVVMKLFGAIGKAYLWESAIHG VVGAVAGSSPAYVEMPEALS 169
QY 184 DGVKMGLEPRRLAIQAGQALLGAARKMLDSEOHPCOLKDNVCSPGATIHAIHFLESQG 243
Db 170 DGAVRGGIPRAQSYEMAAQVLTGAARKMLDSEGKTGALNDWYCSPGGTTIEAVRFLKKG 229
QY 244 FRSLLINAVEASCIRPRELOSM 265
Db 230 MRSSVIEAMTECMESKEFEKI 251

Search completed: January 31, 2002, 14:57:19
Job time: 204 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2002, 14:53:30 ; Search time 23.41 Seconds

(without alignments)
993.550 Million cell updates/sec

Title: US-09-912-717-1
Perfect score: 1567
Sequence: 1 MSVGFIGQLAVRFTAGT.....TPSSPKLLTRSLALGKRD 314

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_1101.*
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1567	100.0	314	21	Human delta 1-pyrr
2	1509	96.3	320	22	Human Py-CR protei
3	1306	83.3	319	22	Human p5CR protei
4	1045	66.7	212	22	Human protein sequ
5	1042	66.5	255	21	Human prostate can
6	537.5	34.3	274	21	Human OXRE-14. Ho
7	537.5	34.3	274	22	Human protein sequ
8	534.5	34.1	274	22	Human dihydropyrr
9	512.5	32.7	276	21	Arabidopsis thalia
10	487.5	31.1	256	21	Arabidopsis thalia
11	402.5	25.7	269	22	Escherichia coli p

12	372	23.7	151	21	AA606980	Arabidopsis thalia
13	370	23.6	270	22	AA690206	C glutamicum prote
14	370	23.6	270	22	AA697982	Corynebacterium gl
15	306.5	19.6	295	22	AA681123	Mycobacterium tube
16	269	17.2	78	21	AA603518	Human secreted pro
17	241.5	15.4	271	22	AA682031	S. epidermidis ope
18	112.5	7.2	92	21	AA634996	Arabidopsis thalia
19	104	6.6	594	14	AA634936	CENP-B. Homo sapi
20	99.5	6.3	341	22	AA690789	C glutamicum prote
21	99.5	6.3	341	22	AA690789	Corynebacterium gl
22	99.5	6.3	341	22	AA690789	Glycerol-3-phospha
23	98	6.3	381	16	AA601433	S. aggregatus PKS
24	98	6.3	1551	21	AA610484	Enzyme involved in
25	97.5	6.2	543	14	AA642456	Biosynthetic enzym
26	97.5	6.2	543	17	AA699465	S. putrefaciens ep
27	97.5	6.2	543	19	AA637053	S. putrefaciens PK
28	97.5	6.2	543	20	AA689403	Shewanella putrefa
29	97.5	6.2	543	21	AA610470	Arabidopsis thalia
30	95.5	6.1	375	21	AA647613	Arabidopsis thalia
31	95.5	6.1	1455	19	AA648663	Arabidopsis thalia
32	94.5	6.0	372	16	AA601432	Glycerol-3-phospha
33	93.5	6.0	393	21	AA616522	Arabidopsis thalia
34	93.5	6.0	393	21	AA647616	Arabidopsis thalia
35	93.5	6.0	542	22	AA664458	Arabidopsis thalia
36	91.5	5.8	1026	18	AA637490	S. putrefaciens ei
37	91.5	5.8	1026	18	AA637490	Caulobacter cresce
38	89	5.7	448	22	AA630237	Corynebacterium gl
39	89	5.7	448	22	AA680046	C glutamicum prote
40	88.5	5.6	349	21	AA647614	Arabidopsis thalia
41	88.5	5.6	408	22	AA691550	Streptococcus pneu
42	88.5	5.6	722	19	AA651502	Streptococcus pneu
43	88.5	5.6	722	19	AA651502	Arabidopsis thalia
44	87.5	5.6	72	21	AA634997	Arabidopsis thalia
45	87	5.6	409	21	AA639651	Arabidopsis thalia

ALIGNMENTS

RESULT 1	AA620584	standard; Protein; 314 AA.
ID	AA620584	
AC	AA620584	
XX		
DT	13-DEC-2000	(first entry)
XX		
DE	Human delta 1-pyrroline-5-carboxylate reductase homologue protein.	
XX		
KW	Human: delta 1-pyrroline-5-carboxylate reductase homologue; P5CRH;	
KW	neurotrophic; neuroprotective; osteopontin; cytoskeletal; antiParkinsonian;	
KW	cardiac; osteopontin; cytoskeletal; hepatotropic; cell proliferation;	
KW	neural; Alzheimer's disease; dementia; Parkinson's disease;	
KW	Alzheimer's disease; dementia; Parkinson's disease;	
KW	Alzheimer's disease; dementia; Parkinson's disease;	
KW	Alzheimer's disease; dementia; Parkinson's disease;	
XX		
OS	Synthetic.	
XX		
PN	US6100075-A.	
XX		
PD	08-AUG-2000.	
PF	18-JUN-1998; 98US-0099676.	
XX		
PR	18-JUN-1998; 98US-0099676.	
XX		
PA	(INCY-) INCYTE PHARM INC.	
XX		
PI	Hillman JL, Baughn MR, Corley NC;	
XX		
DR	WPI: 2000-548292/50.	
DR	N-PSDB: AAA88073.	

XX New human delta 1-pyrroline-5-carboxylate reductase homolog (P5CRH) and
 PT polynucleotides encoding P5CRH, useful for diagnosing, treating or
 PT preventing neuronal or connective tissue disorders, or disorders of
 PT cell proliferation

PS Claim 1: Fig 1A-E; 32pp; English.

XX The present sequence represents human delta 1-pyrroline-5-carboxylate
 CC reductase homologue (P5CRH). P5CRH has neurotropic, neuroprotective,
 CC osteopathic, cytosolic, antiparkinsonian, cardiant, osteopathic,
 CC ophthalmological and hepatotropic activities, and can be used in gene
 CC therapy. P5CRH and the polynucleotides encoding it are useful for
 CC diagnosing, treating or preventing neuronal disorders, connective tissue
 CC disorders, or disorders of cell proliferation. These may be applied to
 CC any subject, e.g. dogs, cats, cows, horses, rabbits, monkeys, or
 CC preferably humans. The polynucleotide sequences may be used to detect
 CC and quantitate gene expression in biopsied tissues where expression of
 CC the polypeptide may correlate with the disease. The diagnostic assay may
 CC be used to determine absence, presence, and excess expression of the
 CC polypeptide, and to monitor regulation of the polypeptide levels during
 CC therapeutic intervention. The polypeptide or its fragments are useful
 CC for treating or preventing neuronal disorder (e.g. akathisia,
 CC Alzheimer's disease, amnesia, dementia or parkinson's disease),
 CC connective tissue disorder (e.g. cystic fibrosis, myocardial fibrosis,
 CC osteoporosis or Marfan syndrome), or a disorder of cell proliferation
 CC (e.g. cirrhosis, leukaemia, lymphoma, sarcoma or bone marrow cancer).
 CC The polypeptide or its fragments are also useful for screening
 CC libraries of compounds in any of several drug-screening techniques.

XX Sequence 314 AA:

Query Match 100.0%; Score 1567; DB 21; Length 314;
 Best Local Similarity 100.0%; Pred. No. 1.9e-159;
 Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVCFGAGQLAYRTAGILSAHKTIASPEMNLPTVSALRKMGNLTRSNKETYKHS 60
 Db 1 msyffigagqlayrtfaagilshahkltasspennlptvsalrkmgnltrsnketykhsd 60
 QY 61 VLFVAVRPHITPFLIDEIGADVQARHIVSCAAGVTISSEVKRLMAFQAPKVIKMTNT 120
 Db 61 vlfavvrphitpflideigadvqarhivscagvtissvekkrlmafqapkvirkmtnt 120
 QY 121 PVMVQEGATVYATGTHALVEDGQLLEQDMSSVGFCTEVEDLIDAVTGISGSPAYAFMA 180
 Db 121 pvmvqegatvyatgthalivegqlleqdmssvgfctevvedlidavtgisgspayafma 180
 QY 181 LDADGVKMGIPRRLAIOIGAQALLGAAMLLDSEOHPCOLKDNVCSPGATTHALHFTLE 240
 Db 181 ldadgvmkgiprrlaioigaqallgaamlldseohpcolkdnvcspgatthalhftle 240
 QY 241 SGGFSRLINAVEASCIRTELQSMADOKETSPALKKTLIDRYKLESPTVSTLTPSSPG 300
 Db 241 sggsfslinaveascirtrelqsmadoketispalkktildryklesptvstltpsspg 300
 QY 301 KLTNSLAIGKKD 314
 Db 301 klttrslalgykkd 314

RESULT 2

AAB74779
 ID AAB74779 standard; Protein; 320 AA.

AC AAB74779;

DT 06-JUN-2001 (first entry)

DE Human Py-CR protein SEQ ID NO:4.

KW Human; pyrroline 5'-carboxylate reductase; Py-CR; P5CR.

XX Homo sapiens.

XX CN1274728-A.

XX 29-NOV-2000.

XX 25-MAY-1999; 99CN-0107071.

XX 25-MAY-1999; 99CN-0107071.

XX (UYFU-) UNIV FUDAN.

XX Yu L, Fu Q, Zhang H;

XX WPI: 2001-211749/22.

XX N-PSDB; AAF81847.

XX New human protein and its code sequence, preparation and application -

XX Claim 4; Page 16; 20pp; Chinese.

XX The present invention describes a human protein designated Py-CR, which
 CC is homologous to the human pyrroline-5'-carboxylate reductase (P5CR)
 CC (EC1.5.1.2). The present invention also describes methods for the
 CC application and production process of the Py-CR polynucleotide and
 CC protein sequences. The present sequence represents the human Py-CR
 CC protein as given in the present invention.

XX Sequence 320 AA:

Query Match 96.3%; Score 1509; DB 22; Length 320;
 Best Local Similarity 96.9%; Pred. No. 3.3e-153;
 Matches 310; Conservative 0; Mismatches 4; Indels 6; Gaps 2;

QY 1 MSVCFGAGQLAYR---FTAGILSAHKTIASPEMNLPTVSALRKMGNLTRSNKETY 56
 Db 1 msyffigagqlanalarqfhaagilshahkltasspennlptvsalrkmgnltrsnkety 60
 QY 57 KHSVDLFAVXPHITPFLIDEIGADVQARHIVSCAAGVTISSEVKRLMAFQAPKVIK 116
 Db 61 khsvdvlfavxphitpflideigadvqarhivscagvtissvekkrlmafqapkvirk 120
 QY 117 MTNPPVYQEGATVYATGTHALVEDGQLLEQDMSSVGFCTEVEDLIDAVTGISGSP 176
 Db 121 mtnppvyqegatvyatgthalivegqlleqdmssvgfctevvedlidavtgisgspay 180
 QY 177 AFMADL--ADGVKMGIPRRLAIOIGAQALLGAAMLLDSEOHPCOLKDNVCSPGATTH 234
 Db 181 afmadaladgvmkgiprrlaioigaqallgaamlldseohpcolkdnvcspgatth 240
 QY 235 ALHFTESGFSRLINAVEASCIRTELQSMADOKETSPALKKTLIDRYKLESPTVSTL 294
 Db 241 alhftesgfsrlinaveascirtrelqsmadoketispalkktildryklesptvstl 300
 QY 295 TPSSPGKLTTRSLAIGKKD 314
 Db 301 tpsspgklttrslalgykkd 320

RESULT 3

AAB74780
 ID AAB74780 standard; Protein; 319 AA.

AC AAB74780;

DT 06-JUN-2001 (first entry)

DE Human P5CR protein.

KW Human; pyrroline 5'-carboxylate reductase; Py-CR; P5CR.

OS Homo sapiens.
 XX CNI274728-A.
 XX 29-NOV-2000.
 XX 25-MAY-1999; 99CN-0107071.
 XX 25-MAY-1999; 99CN-0107071.
 XX (UYFU-) UNIV FUDAN.
 XX Yu L, Fu Q, Zhang H;
 XX MPI; 2001-211749/22.
 DR New human protein and its code sequence, preparation and application -
 XX Example 2; Fig 1; 20pp; Chinese.
 CC The present invention describes a human protein designated Py-CR, which
 CC is homologous to the human pyrroline-5'-carboxylate reductase (P5CR)
 CC (ECL 5.1.2). The present invention also describes methods for the
 CC application and production process of the Py-CR polynucleotide and
 CC protein sequences. The present sequence represents the human P5CR
 CC protein which is used in comparison with the human Py-CR protein in
 CC an example from the present invention.
 XX Sequence 319 AA:
 SQ
 Query Match 83.3%; Score 1306; DB 22; Length 319;
 Best Local Similarity 82.7%; Pred. No. 1, 9e-131;
 Matches 263; Conservative 27; Mismatches 22; Indels 6; Gaps 2;
 QY 1 MSVGTGAGOLAYR---FTAGILSAHKIIASPEENMLPTVALRRKMGVNLTRSKETV 56
 Db 1 msvgtgagolayr-----ftagilslahkiiasspeennlptvalrrkmgvnltrskety 56
 QY 57 KHSVDLFLAVKPHIIPFLDELICADVQARIHVSCAAGVITSSVEKKLMAFOFAPRVINC 116
 Db 61 qhsdvlflavkphiiplfdeidgedicdrhivscagavlitsslekklsalfppaprvinc 116
 QY 117 MTNPPVVOEGATYATGTHALVEDGQLLEQLMSSVGFCTEVEDLIDAVTGSGSGPAY 176
 Db 121 mtnppvvogeatyattgthahvedgqlleqlmssvfgctevbedlidavtgsdsgpay 176
 QY 177 AFMALD--ADGVKMGRLPRRLAIOLGQAOLILGAKKMLDSEQHPCOLKDNVCSFGGATT 234
 Db 181 afmald--adgvkmgrlprrlaioigqaolilgaakmlldseqhpgqlkdnvsspggatt 234
 QY 235 ALHFLSEGGRSLILNAVEASCIPTRELQSMADQEKISPAALKKTLIDRVKLESPVSTL 294
 Db 241 alhvlseggrslllnaveascirrelqsmadqekispaalkktildrvklidspagtal 294
 QY 295 TPSSPGKLLTRSLALGK 312
 Db 301 spsgklllprslapagk 318
 RESULT 4
 ID AAB93027 standard; Protein: 212 AA.
 XX AAB93027;
 XX 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:11795.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KW Homo sapiens.
 OS

XX EPI074617-A2.
 XX 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 XX 27-AUG-1999; 99JP-0300253.
 XX 11-JAN-2000; 2000JP-0118776.
 XX 02-MAY-2000; 2000JP-0183767.
 XX 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isega T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX MPI; 2001-318749/34.
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX Claim 8; SEQ ID 11795; 2537bp + CD ROM; English.
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC of an oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX Sequence 212 AA:
 SQ
 Query Match 66.7%; Score 1045; DB 22; Length 212;
 Best Local Similarity 99.1%; Pred. No. 1e-103;
 Matches 210; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
 QY 105 MAFQAPKVIKMTNTPVVOEGATYATGTHALVEDGQLLEQLMSSVGFCTEVEDLID 164
 Db 1 mafqapkvlrmtntppvvogeatyattgthahvedgqlleqlmssvfgctevbedlid 60
 QY 165 AVTGLSGSGPAYAFMALD--ADGVKMGRLPRRLAIOLGQAOLILGAKKMLDSEQHPCOLK 222
 Db 61 avtglsdsgspayafmald--adgvkmgrlprrlaioigqaolilgaakmlldseqhpgqlk 120
 QY 223 DNVCSFGGATTIHALHFLSEGGRSLILNAVEASCIPTRELQSMADQEKISPAALKKTLID 282
 Db 121 dnvcsfggattihalhflseggrslllnaveascirrelqsmadqekispaalkktilid 180
 QY 283 RYKLESPVSTLTPSSPGKLLTRSLALGKKD 314
 Db 181 rvlklespvstltpsspgkllltrslalggkdd 212

KW Human; dihydropyrrrole-5-carboxylate reductase 30; cancer; cytostatic;
KW human immunodeficiency virus; HIV; infection; immunological disease;
KW inflammatory disease.

OS Homo sapiens.

PN CN1298002-A.

PD 06-JUN-2001.

PF 24-NOV-1999; 99CN-0124090.

PR 24-NOV-1999; 99CN-0124090.

PA (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.

PI Mao Y, Xie Y;

DR MPI: 2001-489680/54.

DR N-PSDB; AAH77597.

PT Human dihydropyrrrole-5-carboxylate reductase 30 as one new kind of
PT polypeptide and polynucleotides encoding this polypeptide -

PS Claim 1; Page 20-21 (disclosure); 26pp; Chinese.

CC The invention relates to a novel polypeptide, human
CC dihydropyrrrole-5-carboxylate reductase 30, polynucleotides encoding
CC this polypeptide and a DNA recombination process to produce the
CC polypeptide. The polypeptide is useful for treating various diseases,
CC such as malignant tumours, nosohaemia, HIV infection, immunological
CC diseases and inflammatory diseases. The invention also provides an
CC antibody against the polypeptide. The present sequence is the
CC polypeptide of the invention.

XX Sequence 274 AA;

SQ

Query Match 34.1%; Score 534.5; DB 22; Length 274;
Best Local Similarity 44.9%; Pred. No. 8.2e-49;
Matches 120; Conservative 46; Mismatches 88; Indels 13; Gaps 5;

OY 3 VGEFGAGC-----LAVRTAGILSAHKIIASSP-EMNLPTVSALRKGVNLTRSNKETVK 57

DB 11 VGIYAGRMAGATAGGIIAGKVEAQHILASAPDNLCHFGAL---GCTHNSQEVILQ 67

OY 58 HSDVFLAVKPHIIPFILDIDIGADVOARHTVSCAAGVTISVYEKKLMAFOPAPKYIRCM 117

DB 68 scIIVfEtKpHvIpaVlaevapvttehlIIVaaqmsIstleelI---pnltrIrvI 124

OY 118 TMTPTVVOEGATVATGTHALVEGQLLEQLMSSVGFCTEVEDLIDAVTGLSGSPAY- 176

DB 125 pIIPcvvgegaIvmargrhvyssetkIlgllleagrcveeprayvdlhtgIsgvatV 184

OY 177 -AFMLDDDGKVMGLPRLAIQLGAQALLGAAKMLDSEQHPCQLKDNVCSFGATTHA 235

DB 185 cefsaaleagvkmgsplahriaqqlIgtkmlIhegnhpqIstvcptggtIy 244

OY 236 LHFLESGGFRSLINAVASCIPTREL 262

DB 245 lhaIegggIraatmsaveaatorakel 271

RESULT 9

AAAG06978

XX AAG06978 standard; Protein; 276 AA.

XX AAG06978;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 3954.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126284.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 04-MAY-1999; 99US-0132048.

PR 05-MAY-1999; 99US-0132407.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132486.

PR 11-MAY-1999; 99US-0132487.

PR 14-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 17-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0149956.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155139.
 PR 24-SEP-1999; 99US-0155486.
 PR 28-SEP-1999; 99US-0155659.
 PR 29-SEP-1999; 99US-0156458.
 PR 04-OCT-1999; 99US-0156596.
 PR 05-OCT-1999; 99US-0157117.
 PR 06-OCT-1999; 99US-0157753.
 PR 07-OCT-1999; 99US-0157865.
 PR 08-OCT-1999; 99US-0158029.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158293.
 PR 13-OCT-1999; 99US-0159294.
 PR 14-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159328.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159638.
 PR 21-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 31.1%; Score 487.5; DB 21; Length 256;
 Best Local Similarity 41.6%; Pred. No. 8.1e-44;
 Matches 107; Conservative 51; Mismatches 92; Indels 7; Gaps 3;

QY 8 AGOLAVFTAGILSAHKITASSPENMLPTVSALRRKGVNLTNRNKETVKHSDVLEFLAVK 67
 Db 2 aesiagvvaagvylpnr-ictavhnlrtdvfaefgvnvtseeevkedvifsvk 60
 QY 68 PHITPILDEIGADVOARHIVSCAGVTIISSEVEKKLAFQAPRVITICMNTPTVVOEG 127
 Db 61 pyvvkavetelkklknkllivsvaagikindlqe----wsgqdfitvmpnptpaavgea 116
 QY 128 ATYVAAGTALHVEDGQLLEBOLMSVSGFCTEVEEDLIDAVTGISGSGPAYARMAID--ADG 185
 Db 117 asvmsigtateedgavlmfegavgkllkadekmfdavtqisgspayitilaaleadq 176
 QY 186 GVMGIPRLAIOLGAGLGAAMKMLLSEOHPCOLKDNVCSPGGATIHALLPLESGGFR 245

Db 177 gvaaglprelalsiasqvlgaatmwsktghkpgvldvotpsggtciagvhelekstr 236
 QY 246 SLINAVEASCIRREL 262
 Db 237 atlmavvaakstr 253

RESULT 11

AAG64106
 ID AAG64106 standard; protein; 269 AA.

AC AAG64106;

DT 21-SEP-2001 (first entry)

DE Escherichia coli proc polypeptide.

KM L-pipecolic acid production; delta-1-piperidine-6-carboxylic acid;
 KM pyrrolidine-5-carboxylate reductase; lysine-6-aminotransferase;
 KW proc.

OS Escherichia coli.

PN WO200148216-A1.

PD 05-JUL-2001.

PF 22-DEC-2000; 2000WO-JP09137.

PR 28-DEC-1999; 99JP-0373389.

PA (SAOC) MERCIAN CORP.

PI Fujii T, Aritoku Y, Mukaihana M, Narita T, Agematu H, Isehiki K,
 DR WPT: 2001-418288/44.

DR N-PSDB; AAH48471.

PT Producing L-pipecolic acid for use as a pharmaceutical raw material,
 PT comprises reducing delta-1-piperidine-6-carboxylic acid with
 PT pyrrolidine-5-carboxylate reductase particularly in recombinant
 PT bacterium

PS Disclosure; Page 31-33; 54pp; Japanese.

CC The invention relates to a method for producing L-pipecolic acid.

CC The method comprises reducing delta-1-piperidine-6-carboxylic acid with
 CC pyrrolidine-5-carboxylate reductase. A recombinant Escherichia coli or
 CC coryneform bacterium that can express a gene encoding

CC lysine-6-aminotransferase is cultured in an L-lysine-containing
 CC medium in order to produce L-pipecolic acid. L-pipecolic acid may be
 CC used as a pharmaceutical raw material. The process is efficient, and
 CC since it is conducted enzymatically, chirality of the final product
 CC can be retained. The present sequence is the Escherichia coli proc
 CC polypeptide. The gene encoding this protein may be used in the
 CC invention.

CC Sequence 269 AA:

Query Match 25.7%; Score 402.5; DB 22; Length 269;
 Best Local Similarity 35.5%; Pred. No. 1.1e-34;
 Matches 98; Conservative 52; Mismatches 107; Indels 19; Gaps 6;

QY 3 VGTIGQALVFTAGILSAHKITASSPENMLPT---VSAIR-KMGVNLTRSNKETVKH 58
 Db 5 igfiggmgng-kallgylasgqvlpqgltwtpdpkvaahdqfinaesqevag 63
 QY 59 SDVLEAVVPHITPILDEIGADVOARHIVSCAGVTIISSEVEKKLAFQAPRVITICMT 118
 Db 64 aditfaavkpgimikvliseltsslnkdsilvsvlaagvldqlar---alghdrklltramp 120

PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	09-JUN-1999	9905-0138540
PR	10-JUN-1999	9905-0138847
PR	11-JUN-1999	9905-0139119
PR	12-JUN-1999	9905-0139452
PR	13-JUN-1999	9905-0139453
PR	14-JUN-1999	9905-0139454
PR	15-JUN-1999	9905-0139454
PR	16-JUN-1999	9905-0139454
PR	17-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139750
PR	18-JUN-1999	9905-0139763
PR	18-JUN-1999	9905-0139817
PR	21-JUN-1999	9905-0139899
PR	22-JUN-1999	9905-0140353
PR	22-JUN-1999	9905-0140354
PR	23-JUN-1999	9905-0140695
PR	24-JUN-1999	9905-0140823
PR	28-JUN-1999	9905-0140991
PR	29-JUN-1999	9905-0141887
PR	30-JUN-1999	9905-0141842
PR	01-JUL-1999	9905-0142154
PR	01-JUL-1999	9905-0142290
PR	02-JUL-1999	9905-0142590
PR	06-JUL-1999	9905-0142803
PR	08-JUL-1999	9905-0142820
PR	09-JUL-1999	9905-0142977
PR	12-JUL-1999	9905-0143342
PR	13-JUL-1999	9905-0143624
PR	14-JUL-1999	9905-0144005
PR	15-JUL-1999	9905-0144085
PR	16-JUL-1999	9905-0144325
PR	19-JUL-1999	9905-0144331
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	19-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144884
PR	21-JUL-1999	9905-0144884
PR	21-JUL-1999	9905-0145086
PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145089
PR	22-JUL-1999	9905-0145129
PR	22-JUL-1999	9905-0145216
PR	23-JUL-1999	9905-0145224
PR	23-JUL-1999	9905-0145276
PR	26-JUL-1999	9905-0145913
PR	27-JUL-1999	9905-0145918
PR	27-JUL-1999	9905-0145919
PR	28-JUL-1999	9905-0145951
PR	02-AUG-1999	9905-0146388
PR	02-AUG-1999	9905-0146389
PR	03-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147560
PR	05-AUG-1999	9905-0147620
PR	06-AUG-1999	9905-0147503

[illegible]

OY 170 SCGSPAYAFMALDA--DGVKMGLEPRRLAIOGAOALLGAAKMLDSEOHPCOLKXNVC 227
 DB 173 sgsspaylflytealltaeaynlgpratakklavaasgeaakmktgkpsellragvas 232
 OY 228 PGGATTIHALHLESGGFRSLINAVEASGCTRTREL 262
 DB 233 pagtvealreleesgrrgalfyragacadrseel 267

RESULT 15

AAG81123
 ID AAG81123 standard; Protein; 295 AA.

XX AAG81123:
 AC
 XX
 DN 04-SEP-2001 (first entry)

DE Mycobacterium tuberculosis potential drug target protein SEQ ID 174.
 XX
 KM Drug target; growth; organism viability; characterisation.
 XX
 OS Mycobacterium tuberculosis.

XX WO200135317-A1.
 PN
 XX
 PD 17-MAY-2001.

PF 13-NOV-2000; 2000WO-US31152.
 XX
 XX

PR 12-NOV-1999; 99US-0165086.
 PR 12-NOV-1999; 99US-0165124.
 PR 01-FEB-2000; 2000US-0179531.

XX (REGC) UNITV CALIFORNIA.
 PA
 XX

PI Eisenberg D, Rotstein SH, Marcotte EM;
 XX
 DR WPI: 2001-329193/34.
 DR N-PSDB; AAH51974.

XX
 PT Identifying nucleotide or polypeptide sequence for use as drug target,
 PT involves providing algorithm that analyzes a functional relationship
 PT between nucleotide or polypeptide sequences, and comparing the
 PT sequences .

XX
 PS Disclosure; Page 161; 207pp; English.

XX
 CC This invention relates to a method for identifying a nucleotide or
 CC polypeptide sequence that may be a drug target, or essential for growth
 CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
 CC represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium
 CC tuberculosis proteins which are potential drug targets. The DNA and
 CC protein sequences are used to illustrate the method of the invention. The
 CC method involves providing an unknown nucleotide or polypeptide sequence,
 CC and comparing it to a number of sequences along with at least one
 CC algorithm capable of analysing a functional relationship between
 CC nucleotide and polypeptide sequences. The method is useful for
 CC characterising the function of nucleic acids and polypeptides that may be
 CC useful as a target for a drug or essential for the growth or viability of
 CC an organism.
 XX

XX
 SK Sequence 295 AA;

Query Match 19 6%; Score 306.5; DB 22; Length 295;
 Best Local Similarity 34.3%; Pred. No. 2.6e-24;

Matches 86; Conservative 39; Mismatches 91; Indels 35; Gaps 6;

OY 48 LTRSNKEFYKHSVDVLEAVKP-HIIPFLDEGADV-----QAKHIYSCAGVTISSVE 101
 DB 54 lvsaaadaventalftvavkpadvprvialanaaanaaendsaegvftvavgtllyafe 113

OY 102 KRLAFOAPAKVIRCTNTPTVVYVEGATVYATGTHALVEDGOLLBOLMSSVGEFCTEVED 161
 DB 114 sklpaqcp---vvvrampnaaalvgaytalakgrfvtppqlaevsalldavgvltvres 170
 OY 162 LIDAVTGLSGSPRAYAFMALDA--DGVKMGLEPRRLAIOGAOALLGAAKMLDSEOH-- 217
 DB 171 qldavtavsgspayfflllvealvdagvgyglrvaldaaqtmagsaamltermedq 230
 OY 218 -----PCOLKDNVCSPGAT IHALHLESGGFRSLINAVEASGCTRTREL 262
 DB 231 ggangetlmglrvalltasriraavtspgltlaaalreleergfrmeavdaavga----- 283
 OY 263 QSMADQEKISP 273
 DB 284 ksrseqrlrltp 294

Search completed: January 31, 2002, 14:54:44
 Job time: 74 sec

Thu Jan 31 15:07:53 2002

us-09-912-717-1.rag
